

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 19, 2002, 02:23:32 ; Search time 2737 Seconds

(without alignments)
14758.748 Million cell updates/sec

Title:	US-09-905-674-1
Perfect score:	1 298

Sequence: 1 ctctctgcccagccggc.....gggagccnagacgttqccc 1388

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 2000000000

Post-processing:	Minimum Match	0%
Maximum Match	100%	

Listing first 45 summaries

Database :

```

1:  Genemb1.*
2:  1:  gb_ba.*
3:  2:  gb_htg.*
4:  3:  gb_in.*
5:  4:  gb_om.*
6:  5:  gb_ov.*
7:  6:  gb_pat.*
8:  7:  gb_ph.*
9:  8:  gb_pl.*
10:  9:  gb_pr.*
11:  10:  gb_ro.*
12:  11:  gb_sls.*
13:  12:  gb_sy.*
14:  13:  gb_un.*
15:  14:  gb_vl.*
16:  15:  gb_ba.*
17:  16:  em_fun.*
18:  17:  em_hum.*
19:  18:  em_in.*
20:  19:  em_mu.*
21:  20:  em_om.*
22:  21:  em_or.*
23:  22:  em_ov.*
24:  23:  em_pat.*
25:  24:  em_ph.*
26:  25:  em_pl.*
27:  26:  em_ro.*
28:  27:  em_sls.*
29:  28:  em_un.*
30:  29:  em_vl.*
31:  30:  em_htg_hum.*
32:  31:  em_htg_in.*
33:  32:  em_htg_other.*
34:  33:  em_htg_mu.*
35:  34:  em_htg_pl.*
36:  35:  em_htg_ro.*
37:  36:  em_htg_sls.*
38:  37:  em_htg_sy.*
39:  38:  em_htg_un.*
40:  39:  em_htg_vl.*
41:  40:  em_htg_om.*
42:  41:  em_htg_ov.*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1385.2	99.8	1388	6	AX440923	AX440923 Sequence
2	1332.2	96.0	2556	9	HSB801611	AL136638 Homo sapi
3	1323.4	95.3	2553	9	AF311903	AF311903 Homo sapi
4	1214	87.5	2502	9	BC0002920	BC0002920 Homo sapi
5	816.2	58.8	2428	10	BC0255568	BC0255568 Homo sapi
6	815.2	58.7	2498	10	BC024611	BC024611 Homo sapi
7	811.8	58.5	2500	10	BC026574	BC026574 Homo sapi
8	660	47.6	2605	9	AK055530	AK055530 Homo sapi
9	541.8	39.0	161804	9	AC021028	AC021028 Homo sapi
10	362	26.1	813	6	AX420466	AX420466 Sequence
11	362	26.1	3184	6	AX420466	AX420466 Sequence
12	348.8	25.1	1516	10	BC010366	BC010366 Sequence
13	345.2	24.9	2426	9	BC010405	BC010405 Sequence
14	340.2	24.5	3175	10	AF121344	AF121344 Homo sapi
15	337.4	24.3	1405	9	AF065389	AF065389 Homo sapi
16	337.4	24.3	1416	9	BC009704	BC009704 Homo sapi
17	303	21.8	1408	9	AF055455	AF055455 Homo sapi
18	230	16.6	864	6	AX343015	AX343015 Sequence
19	216.4	15.6	1567	6	AX061778	AX061778 Sequence
20	206.4	14.9	557	6	AX079573	AX079573 Sequence
21	203	14.6	4445	6	AX024427	AX024427 Homo sapi
22	178	12.8	852	6	AX247836	AX247836 Sequence
23	177.6	12.8	1995	10	BC024685	BC024685 Mus muscu
24	158.4	11.4	179772	2	AC108430	AC108430 Mus muscu
25	135.2	9.7	143717	2	AC123389	AC123389 Rattus no
26	134.8	9.7	143717	2	AC123389	AC123389 Rattus no
27	134.8	9.7	167399	2	AC103212	AC103212 Rattus no
28	130.4	9.4	1216	3	AF274013	AF274013 Drosophi
29	124.4	9.0	1703	9	AF120266	AF120266 Homo sapi
30	124.4	9.0	1726	9	BC003157	BC003157 Homo sapi
31	124.4	9.0	1734	9	BC004161	BC004161 Homo sapi
32	123.8	8.8	1778	6	AX376210	AX376210 Sequence
33	122.4	8.8	1661	10	BC003872	BC003872 Mus muscu
34	115.6	8.2	1601	9	AF125213	AF125213 Homo sapi
35	113.4	8.2	368	9	AF174603	AF174603 Homo sapi
36	112.8	8.1	1693	9	BC032802	BC032802 Homo sapi
37	111.6	8.0	1646	10	BC019991	BC019991 Mus muscu
38	105	7.6	167399	2	AC103212	AC103212 Rattus no
39	95.4	6.9	394	6	AX072723	AX072723 Sequence
40	91.8	6.6	389	6	AX072711	AX072711 Sequence
41	91.8	6.6	102468	6	AC000994	AC000994 Drosophi
42	91.8	6.6	106584	2	AC013977	AC013977 Drosophi
43	91.8	6.6	140973	3	AC008140	AC008140 Drosophi
44	91.8	6.6	159007	3	AC009219	AC009219 Drosophi
45	91.8	6.6	224795	3	AE003668	AE003668 Drosophi

ALIGNMENTS

RESULT 1				
AX440923				
LOCUS	AX440923	1388 bp	DNA	Linear
DEFINITION	Sequence 1 from Patent WO0206340.			PAT 28-JUN-2002
ACCESSION	AX440923			
VERSION	AX440923.1	GI:21665563		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Reinhard,C.J. and Garcia,P.D.			
TITLE	Teitraspan protein and uses thereof			
JOURNAL	Patent: WO 0206340-A 1 24-JUN-2002;			


```

/dB_xref="GI:22266722"
/translation="MHYRYISNAKVCSCWKYKLFESYNIIFPLAVVFLGCVGLAMNSEK
GLVSLDKTRTMRMGHIDPVPVLVLMGVAMFTLFGAGVGALENICLINFPGTIVLH
FLELAVAVLAFLEQDWWDRREFRESNISKYSRDDIDLSLOKAMDCGAGYGP
DMIDLVAVENCSCASREKCGPFCSCPDACVAVDVTQCYDVRIOLKSKMSSEIF
KGCIOALESMIPRIYIVAGVFIAISLQIGIFLARTLSDIEAVKAGHHF"
BASE COUNT      505 a      724 c      680 g      644 t
ORIGIN

```

[illegible]

QY	901	AGCTGAGCCAGCAGCTGGAGGGCCAGAGCCCTTTCTGCAATCAGCCCTCAGTCCAGAGGA	960
Db	907	AGCTAGGCCAGCGTGGAGGGCCAGAGCCCTTTCTCTCCATCAGCCCTCAGTCCAGAGGA	966
QY	961	GAGGAGCCGACACCCCCAGAGGCCAGTGGCCCATCTTAAGCATCAGCTGACGTACCTCT	1020
Db	967	GAGGAGCCGACACCCCCAGAGGCCAGTGGCCCATCTTAAGCATCAGCTGACGTACCTCT	1026
QY	1021	CTGTTTCTGCTTGCTGCTGCTGTAACCAAGGTCGCCCTTTACGTCGCCAAACTGT	1080
Db	1027	CTGTTTCTGCTTGCTGCTGCTGTAACCAAGGTCGCCCTTTATCCTGTCGCCAAACTGT	1086
QY	1081	GAGCTCATCCCTCTGAGAGTCTACCCAGAGACAGAGAACTGTCTTTATGTGGAGTGTGT	1140
Db	1087	GAGCTCATCCCTCTGAGAGTCTACCCAGAGACAGAGAACTGTCTTTATGTGGAGTGTGT	1146
QY	1141	ACTGTGAAGACAGAGAGGGCTCTGTGTCTCCAGAGAGGCTTGAATCAGACCCCTGTC	1200
Db	1147	ACTGTGAAGACAGAGAGGGCTCTGTGTCTCCAGAGAGGCTTGAATCAGTGA--CCCTGTCC	1205
QY	1201	ACCTCAGAGCATGTGACAGACACCCCTGTGCCCTCTCTCAATTGGCAWTCACACATCTTGC	1260
Db	1206	ACCTCAGAGCATGTGACAGACACCCCTGTGT--CCCTCTCACACTGGGATCCAGACATCTGC	1264
QY	1261	TTTGGGTATCCACATCTGTGGGTGAGCCGCTGGGTAGAGGACCCACAGAGGCTGACAGG	1320
Db	1265	TTTGGGTATCCACATCTGTGGGCGGGCCGTGGGTAGAGGAGCCACACAGTGTGGACAGG	1324
QY	1321	GGATCTCTTCATCAAGCAAGACAGCATGGGGGGCTTCCGTTAAACGGAGGGG	1376
Db	1325	GGATCTCTTCATCAAGCAAGACAGCATGGGGG--CTGCCCCGTAAAGGGAGGGG	1378

RESULT 4
 LOCUS BC002920
 DEFINITION
 2502 bp mRNA linear Pfl 12-JUL-2001
 Homo sapiens, Similar to transmembrane 4 superfamily member 9,
 clone MGC:11352 IMAGE:3554042, mRNA, complete cds.
 ACCESSION BC002920
 VERSION BC002920.1 GI:12804132
 KEYWORDS
 SOURCE MGC.
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2502)
 Authors Strausberg, R.
 Title Direct Submission
 Journal Submitted (05-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 2498)
JOURNAL Strausberg, R.
Direct Submission
Submitted (01-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarline, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 37 Row: e Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction, Similarity but not identity to protein.
FEATURES
Source
1..2498
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:28503 IMAGE:418261"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_S62"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
55..867
/codon_start=-1
/product="Similar to region containing hypothetical
protein MGC11352; slingshot"
/protein_id="AAH24611.1"
/db_xref="GI:19354054"
/translation="MRTYRISNAEWSCKRYLLSFYNIVFWLAGVFLGVGIMANSEK
GVLSDLKTVRLHGDIPVVLVLMVGVMTTLGFAGCVGLARENITLKFPGCAVLIF
FLBLAAVLAFLFQDWRDREFPESNLSYRDKIDIDLNLIDSLQKANOCCGAATPE
DMDLVNFNCSGASYSREKGVFPSCVDPKQVNVNCGDVAIOLKSKDEFIFT
KCIOLAEGLPNINIVAGVFALSLDIFGIFLARFLISDIEAVKGNHP"
CDS
BASE COUNT 542 a 644 c 631 g 681 t
ORIGIN
Query Match 58.7%; Score 815.2; DB 10; Length 2498;
Best Local Similarity 83.4%; Pred. No. 1,6e-176;
Matches 1013; Conservative 0; Mismatches 183; Indels 18; Gaps 7;
13 AGCCGGCGCGCGCGCTGCCCGCGCGCGCGATCTGCTTCTCGAGAGATGCAC 72
Db 1 AACCGGGCGCTGCGCGCGCGCTGCTGCCCGCGCGCGCTCTCTCTCGAGAGATGCAC 60
Qy 73 TATATAGATACCTTAAGCCCAAGTCACTGCTGTACAGTACCTCTTTTCAGCTAC 132
Db 61 TATATAGATACCTGAAGCGCGAGTCACTGCTGTACAGTACCTCTCTCTCACTAC 120
Qy 133 AACATCATCTTCTGTTGGCTGTGAGTTGTCTCTTGGAGTCGCGGCTGTGGCATGGAG 192
Db 121 AATATCGCTTTTGGCTGTGAGTTGTCTCTTGGAGTTCGGGCTGTGGCATGGAGC 180
Qy 193 GAAAAGGTGTGTCTCCGACCTCAACCAAGTGAACCCGGATGATGAGATGACCTGTG 252

Db 181 GAAAAGGTGTGTCTCCGACCTCAACCAAGTGAACCCGGATGATGAGATGACCCCGCTC 240
Qy 253 GTGTGTGTCTGATGTGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 312
Db 241 GT 300
Qy 313 GCTGT 372
Db 301 GCCCTCCGAGAAACATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Qy 373 TTCTGTGAGCTGT 432
Db 361 TTCTGTGAGCTGT 420
Qy 433 TTCTGTGAGCTGT 492
Db 421 TTCTGTGAGCTGT 480
Qy 493 CTGATGACCTCTTCAAGAAAGCTTAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 552
Db 481 CTGATGACCTCTTCAAGAAAGCTTAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 553 GACCTCAAGCTCTTCAAGT 612
Db 541 GACCTCAAGCTCTTCAAGT 600
Qy 613 CCTCTGT 672
Db 601 CCTCTGT 660
Qy 673 GATGTGAGT 732
Db 661 GATGTGAGT 720
Qy 733 CAGCGCTGTGAAAGCTGT 792
Db 721 CAGCGCTGTGAAAGCTGT 780
Qy 793 ATCTGT 852
Db 781 ATCTGT 840
Qy 853 GCAGTGAAGGCGCGCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 912
Db 841 GCAGTGAAGGCGCGCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899
Qy 913 CTGT 972
Db 900 TGT 953
Qy 973 CCCCAGAGCGCGCGCATCTTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
Db 954 CACCAGAGCGCGCGCATCTTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1013
Qy 1033 GCTGT 1092
Db 1014 GCTGT 1070
Qy 1093 -CTGT 1150
Db 1071 GAAAAGGCTGT 1126
Qy 1151 ACAGAGAGGCTGT 1210
Db 1127 GAGAGAGGCTGT 1184
Qy 1211 TGCTGTGAGACAC 1224
Db 1185 CATCTGCAAGGCTC 1198

RESULT 7

OY	777	TGSGCTTCATCGCCATCTCCTGTTCAGAGATTTTGGCATCTTCTTGCAAGAAGCGT	836
Dd	110885	TGSGCGCTGACTCGCTGTGTGTGTTTCAGATATTATTGGCATCTTCTTGCAAGAAGCGT	110944
OY	837	GATCTCAGACATCGAAGGCACGTGAAGGCCGGCCATCTCTGAGAGACAGATTTGAGGGA	896
Dd	110945	GATCTCAGACATCGAAGGCACGTGAAGGCCGGCCATCTCTGAGAGACAGATTTGAGGGA	111004
OY	897	GCCGAGCTGAGCCACGCTGGGAGGCCAGACCTTTCTCTGCCATCAGCCCTACGTGCAGA	956
Dd	111005	GCCGAGCTGAGCCACGCTGGGAGGCCAGACCTTTCTCTGCCATCAGCCCTACGTGCAGA	111064
OY	957	GGGAGAGAGAGCCGACACCCCCAGAGCCAGTGCOCATCTTAAGCATCAGCGTAGCTGAC	1016
Dd	111065	GGGAGAGAGAGCCGACACCCCCAGAGCCAGTGCOCATCTTAAGCATCAGCGTAGCTGAC	111124
OY	1017	CTCTCTGTTTTGCTGTGCTGTGCTGATAAGAACAAAGGCTCCCTGTGATCCGTGCCCAAC	1076
Dd	111125	CTCTCTGTTTTGCTGTGCTGTGCTGATAAGAACAAAGGCTCCCTGTGATCCGTGCCCAAC	111184
OY	1077	TTGTGACTGACATCCCTCTGAGAGTGTACCACAGAGACAGACAAGATGTGHTTATGTGGAGT	1136
Dd	111185	TTGTGACTGACATCCCTCTGAGAGTGTACCACAGAGACAGACAAGATGTGHTTATGTGGAGT	111244
OY	1137	GGTGACTCTGAAAGACAGAGAGGGCTCCTGTGCTGCCAGAGAGGCTTGACTAGACCCC	1196
Dd	111245	GGTGACTCTGAAAGACAGAGAGGGCTCCTGTGCTGCCAGAGAGGCTTGACTAGACCCC	111304
OY	1197	CTGCAAGCTCAAGCATGTGTGAGAGACACCGTGGTCCCTCTCCATYGAGGWTCCAGACAT	1256
Dd	111305	CTGCAAGCTCAAGCATGTGTGAGAGACACCGTGGTCCCTCTCCATYGAGGWTCCAGACAT	111363
OY	1257	CTGCTTTGGGTCATCCACATCTGTGGTNGGCCCTGTGGGTAGAGGAGCCACAGCGTGA	1316
Dd	111364	CTGCTTTGGGTCATCCACATCTGTGGGNGGCCCTGTGGGTAGAGGAGCCACAGCGTGA	111423
OY	1317	CAGGSCATCTCTCTCCATCAAAGCAAGCAGATGGGGGGCTTGCCGTAACGGAGAGCG	1376
Dd	111424	CAGGSCATCTCTCTCCATCAAAGCAAGCAGATGGGG--CTGCGCGGTATACGGAGAGCG	111481

	RESULT	10					
AX420468	LOCUS	AX420468	813 bp	DNA	linear	PAT 18-JUN-2002	
	DEFINITION	Sequence 3 from Patent WO0216603.					
	ACCESSION	AX420468					
	VERSION	AX420468.1	GI:21524616				
	KEYWORDS						
SOURCE							
ORGANISM		human.					
		Homo sapiens					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
REFERENCE							
AUTHORS		1 Leiby, K.R.					
TITLE		23228, a human tetraspanin family member and uses thereof					
JOURNAL		Patent: WO 0216603-A 3 28-FEB-2002;					
		Millennium Pharmaceuticals, Inc. (US)					
FEATURES		Location/Qualifiers					
source		1..813					
		/organism="Homo sapiens"					
		/db_xref="taxon:9606"					
BASE COUNT		151 a 233 c 235 g 194 t					
ORIGIN							

Query Match	26.1%	Score 362;	DB 6;	Length 813;
Best Local Similarity	66.9%;	Pred No. 1.6e-72;		
Matches 515; Conservative	0;	Mismatches 255;	Indels 0;	Gaps 0;
<hr/>				
OY	95 AAGTCACGCTCGTGTAACAAGTACCCTCCCTTTTACGTACAACAATCAACTTCTGGTTGGCTG	154		
Db	35 AGGTGCCTCCTCGGAATAAATTCTCTGTTGGCTTCAACAATGCTTCTGGGTCTGG	94		

OY	155	GAGTTTGTCTTCCTTGAGATCGGGCTGAGGCGATGAGAGCGAAAGGTTGCTGTCGACG	214
Db	95	GAGCCCTCTTCTCGGCTATCGGCTCTGGGGCTGGGGGTAGAGAAAGCGCTTCTCTGAAACA	154
OY	215	TCACCAAGTGCACCCGAGTGCATGGAATCGACCCCTGTGTGCTGGTCCCTGATGCTGGGCG	274
Db	155	TCTCAGCCCTGACAGATGTGGGAGGCGCTTGACCCCGTGTGCTGTTTGTGTAGTTGGAG	214
OY	275	TGTGATGTTTCAACCTGGGGGTTCCGGGCTCGTGGGGGCTCTCGGGAAGATATCTGCT	334
Db	215	GCGTATGTCGAGTGTGGGCTTGTGCTGCTCATTTGGGGGCGCTCCGGAGAACCTTTCC	274
OY	335	TGCTCACTTTTTCGTGTGGACACATCGTCTCATCTTTCTCTGGAGCTGGCTGTGGCCG	394
Db	275	TGCTCAAGTTTTTCTTCGCTGTTCCCGGCTCTCATCTTCTCTGGAGCTGGCAACAGGA	334
OY	395	TGCTGGCCCTTCTCTGTTTCAGGACTGGGTGAGGAGACCGGTTCCGGGAGTTCTTTCGAGACA	454
Db	335	TCTCTGCGCTTGTCTTCAAGGACTGTGATTCGAGACACACTCAACCTCTTCTATCAACAACA	394
OY	455	ACATGAATGCTACCGGGAGACGATTCGATTCGCAAAACCTCATGACTCCCTTGAGAAAG	514
Db	395	AGCTCAAGGCCCTACCGGGAGGACATTTGACTCTCCAGAACCTCATTTGCTGCTAGGAAT	454
OY	515	CTAACCACTGCTGTGGCGCATATGGCCCTGGAAGACTGGAGCCTCAACGTTCACTTCAATT	574
Db	455	ACTGGCTTGTGTTGAGAGCCGAGCCCAATGACTGAACTCAATATCTCACTCAACT	514
OY	575	GGAGCGGTGCCAGCTACAGACCGGAGAAAGTGGGGGGTCCCTTCTCGTGCCTGGCCAG	634
Db	515	GCACTGACTTGAACCCCAACCCCGGAGCGCTGCGGGGTGCCCTTCTCTGCTGCTCAAGG	574
OY	635	ATTCCTGCGAAAAAGTTGTGAACACACAGTGTGATATGATGTCAAGATTCACTGAAAGA	694
Db	575	ACCCGTGCGAGAGATGTCCTCAACACCCAGTGTGGCTAGAGACTCCGGCTCAAACTGGAGC	634
OY	695	GCAAGTGGGATGAGTCCATCTTCACGAAAGGTGCATCCAGCGCTGGAAAGCTGGCTCC	754
Db	635	TGAGAGAGAGGAGCTTCATCCACACCAAAAGGTGGGTGGCCAGTTTGAAGAGTGGCTGC	694
OY	755	CGGGAACATTTACATTGTGGCTGGCGTTTCATCGCCATCTCGCTTGTGACGATATTTCG	814
Db	695	AGGACAACCTGATTTGTGTGGGAGGTCTTCATGGGCTATCCCGCTCCAGATCTTTTG	754
OY	815	GCATCTTCTCGGCAAGAGACGTGATCTCAGACATCGAAGGCACTGGAAGGCC	864
Db	755	GCATCTGCTGGCCAGAACCTCTGTAGTGCATCAAGCACTGGAAGGCC	804

RESULT 11					
AX420466					
LOCUS	AX420466	3184 bp	DNA	linear	PAT 18-JUN-2002
DEFINITION	Sequence 1 from Patent WO0216603.				
ACCESSION	AX420466				
VERSION	AX420466.1	GI:21524614			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1				
AUTHORS	Leiby, K.R.				
TITLE	23228, a human tetraspanin family member and uses thereof				
JOURNAL	Patent: WO 0216603-A1 28-FEB-2002;				
	Millennium Pharmaceuticals, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..3184				

```

CDS      168..980
         /note="unnamed protein product"
         /codon_start=1
         /protein_id="CAD35489.1"

```

```

/db_xref="GI:21524615"
/translation="MPGKHQHFQDEPEVCCGKRYFLGFENVIVWLGALFLAIGLWANG
EKGVLSNISALDGLDIPVWLFVYGVASVYLGAGCIGALRENTPLIKFSPVFLG
IFLELATGILAFVFDWIRDOJLNFINNVAAYRDIDLOJLIDFADWYMSCCGARG
PNDMNLIYFNCNDLNPBSRRCGVPSCVRPAEDVLTGCGYDIRUKLELEDOGSI
HTKGCVOEFKRWLODNLIVAGLVGIALLOJFGLICLQNLVSDIKAVKANM"
BASE COUNT      675 a      883 c      945 g      681 t
ORIGIN
Query Match      26.1%; Score 362; DB 6; Length 3184;
Best Local Similarity 66.9%; Pred. No. 1.6e-72;
Matches 515; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY  95 AAGTACGTGCTGTACAGTACTCTTTACAGTCAACATCATCTTGTGCTG 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  202 AGGTGCGCTGTGGGAAATATCTCTGTTGCTTCAACATGTCCTGGGCTG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  155 GAGTTGCTTCTTGAGATCGGGCTGTGGCATGAGCAAAAGGCTGCTGCGAGC 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  262 GAGCCCTTCTGCTGCTATCGGCTCTGGGCTGTGGATGAGMAGGCTTCTCGACA 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  215 TCACCAAGTACCCGGATGATGCAATCGACCTGTGTGCTGCTGATGCTGGGCG 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  322 TCTCAGCCCTGACAGATCTGGGAGCCTTGACCCCTGTGGCTGTTGTGTAGTGGAG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  275 TGGTATGTTGACCCCTGGGCTGGCGGCTGCGGAGGCTGCGGAGAAATCTGCT 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  382 GCGTATGTCGGTGTGGCTGTTGCTGTGCTGATGGGCGCTTCGGAGAACACCTTCC 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  335 TGCTCAACCTTTTCTGTGGACACATGCTCATCTTCTCTGAGACCTGCTGAGCCG 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  442 TGCTCAACCTTTTCTGTGGACACATGCTCATCTTCTCTGAGACCTGCTGAGCCG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  395 TGCTGAGCTTCTGTGTCAGAGTGGGAGGAGGAGGCTTCCGGAGTCTTGAAGCA 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  502 TCTGAGCTTGTGTCTTCAAGAGTGTGATGAGACACAGCTCAACCTTCTCATCAACA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  455 ACATCAAGTCTTACCGGAGATATGATGCAAAACCTCATGACCTCCCTTCAGAAAG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  562 AGCTCAAGGCTTACCGGAGATATGATGCAAAACCTCATGACCTCCCTTCAGAAAG 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  515 CTAACAGTGTGTGGGAGATATGAGCTTGAAGCTGAGACCTCAAGCTTCAATT 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  622 ACTGTCTGTGTGTGAGACCCGAGGCCCAATGACTGAACTCAATATCTTCAACT 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  575 GCAGCGGTGCGAGCTACAGCCGAGAGAGTGGGGGTCCCTTCTCTGCTGCTGCGAG 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  682 GCAGTGAAGTGAACCCCGAGGAGGAGGCTGCGGGGTCCCTTCTCTGCTGCTGAGG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  635 ATCTGCGCAAAAGTGTGAACACAGTGTGATATGATGTGAGATTCAAGTGAAGA 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  742 ACCCTGCGGAGATGTCTCTCAACACCCAGTGTGCTGACGAGCTCGGCTCAAACTGAGC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  695 GCAAGTGGGATGAGTCCATCTTCAAGAAAGGCTGATCCAGGCGCTGGAAGGCTGTC 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  802 TGGAGCAGCAGGAGCTTCAATCCACACCAAGGCTGCTGGGCAAGTTGAGAAAGTGGCTGC 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  755 CGCGAACAATTTTACATTTGTGCTGTGGCTCTTCAATCGCATCTGCTTGCAGATATTG 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  862 AGGACAACCTGATTTGTGTGGGAGGAGTCTCATGGGCAATCGCCCTCTCAGATCTTTG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  815 GCATCTCTGCGAAGGAGCTGATCTGACATCGAGGAGGAGTGAAGGCC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  922 GCATCTGCTGGCCGAGAACCTGCTGAGTGAATCAAGGAGTGAAGGCC 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

KEYWORDS
SOURCE      house mouse.
ORGANISM    Mus musculus
TITLE       Mammalia: Eutheria: Rodentia: Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1516)
AUTHORS     Strausberg, R.
JOURNAL     Direct Submission
SUBMITTED   (05-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
CONTACT     MGC help desk
EMAIL       cgabs-f@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/HLN at: http://Image.Hlnl.gov
Series: IRK Plate: 5 Row: f Column: 18.
Location/Qualifiers
1. 1516
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="C57BL/6J"
/clone="MGC:6941 IMAGE:2811935"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_id="NCI-CGAP_Mam5"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
136..948
/codon_start=1
/product="RIKEN cDNA 2210021G21 gene"
/protein_id="AAH10346.1"
/db_xref="GI:16307593"
/db_xref="LocusID:74257"
/translation="MPGKHQHFQDEPEVCCGKRYFLGFENVIVWLGALFLAIGLWANG
EKGVLSNISALDGLDIPVWLFVYGVASVYLGAGCIGALRENTPLIKFSPVFLG
IFLELATGILAFVFDWIRDOJLNFINNVAAYRDIDLOJLIDFADWYMSCCGARG
PNDMNLIYFNCNDLNPBSRRCGVPSCVRPAEDVLTGCGYDIRUKLELEDOGSI
HTKGCVOEFKRWLODNLIVAGLVGIALLOJFGLICLQNLVSDIKAVKANM"
BASE COUNT      277 a      417 c      461 g      361 t
ORIGIN
Query Match      25.1%; Score 348.8; DB 10; Length 1516;
Best Local Similarity 65.6%; Pred. No. 1.7e-69;
Matches 509; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY  89 ACGCAAGTACAGCTGCTGACAGTACCTCTTTCACATCAACATCATCTTGTGAT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  164 ACCCTGAGAGTGTGCTGCTGCGGAGAAATATCTTGTGCTTCAACATGCTTCTGGG 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  149 TGGCTGAGTGTCTTCTTGAAGTGGGCTGTGGCATGAGAGGAAAGGCTGCTG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  224 TGGCTGAGAGTGTCTTCTTGAAGTGGGCTGTGGCATGAGAGGAAAGGCTGCTG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  209 CCGACCTCACCAAGTACCCGAGTGCATGAGATCGACCTGTGCTGCTGATGAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  284 CCAATATCTGCGGCTGACATCTGCGGAGTGTGACCCGCTGTGCTGCTGCTG 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  269 TGGCGGTGATGTTCAACCTGGGAGTTCGCGGCTGCTGCGGAGTGAATA 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 755 AGGACACCTGATTTGGTGGCGGAGTCTTATGGGATCGCCCTCCAGATCTTG 814
QY 815 GCATCTTCTGGCAAGAGCGCT 836
Db 815 GCATCTGCTGGCCAGACCT 836

RESULT 14
AF121344 3175 bp mRNA linear ROD 01-FEB-2000
LOCUS Mus musculus tetraspanin Tspan-5 (Tspan5) mRNA, complete cds.
DEFINITION AF121344
ACCESSION AF121344
VERSION AF121344.1 GI:6841032
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3175)
AUTHORS Garcia-Frigola,C., de Lecea,L. and Soriano,E.
TITLE Mouse Tspan-5 cDNA cloning
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3175)
AUTHORS Garcia-Frigola,C., de Lecea,L. and Soriano,E.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1999) Dept. of Animal and Plant Cell Biology,
University of Barcelona, Av. Diagonal 645, Barcelona 08028, Spain

FEATURES
source
1..3175
/organism="Mus musculus"
/strain="Swiss Webster/NIH"
/db_xref="taxon:10090"
1..3175
/gene="Tspan5"
388..1194
/gene="Tspan5"
/note="member of transmembrane 4 superfamily"
/codon_start=1
/product="tetraspanin Tspan-5"
/protein_id="AA028869.1"
/db_xref="GI:6841033"
/translation="MSGKHYKGPVSCIKYIFGFENVIFWFLGITFLGILMANEK
GVLNLSITDLDGDPVWFLVGVGMFLIFGACIALRENTFLKFFSVFLGILF
FLFIAGVLAFFVKWIKDQLEFIFINNIRAYDDIDLONLIDFQEWKCGAGFAD
DMLNLYFNCTDSNASRERKGVPSCTPDPEVDINOCGDAROKREVOQOYIYIT
KGCVPQPEKWLQDNLITIVGIFIGIALQIFICLAQNLVSDIEAVRASW"

BASE COUNT 667 a 811 c 840 g 857 t

ORIGIN
Query Match 24.5%; Score 340.2; DB 10; Length 3175;
Best Local Similarity 63.8%; Pred. No. 1.6e-67;
Matches 516; Conservative 0; Mismatches 293; Indels 0; Gaps 0;

QY 59 CTCGAGAGGACGACTATTATAGATCTACAGCCCAAGGACGCTGCTCAAGTACC 118
Db 380 CCCAGAGAGATGTCGGAGACACTACAGGGCTCTGAAGTCAGTTTTCATCAATACT 439
QY 119 TCCCTTTCAGCTACACATCATCTTCTGGTGGAGTGTCTTCTTGGAGTGGGC 178
Db 440 TCATTTTGGCTCAATGTCATATTTGGTTTGGAAATACGTTCTTGGAAATCGGAC 499
QY 179 TGTGGGATGGAGGAAAGGGTGTGTCGACCTCCACCAAGTACCGGATGCATG 238
Db 500 TGTGGGATGGAGGAAAGGGTGTGTCGACCTCCACCAAGTACCGGATGCATG 559
QY 239 GAATGACCTGTGTGCTGCTGCTGATGTGGGCGGTGATGTTACCTGGGGTTG 298
Db 560 GCTTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 619
QY 299 CCGGCTGCTGGGGCTCTGGGAGAGATATGCTTGCCTCACTTTTCTGAGGACCA 358
Db 620 CAGGTGATGAGGACCTGGGAGAAACACTTCTTCTCAAGTTTCTTCTGCTGTC 679

QY 359 TCGTGTCACTTCTTCTCTGAGAGCTGCTGTCGCTGCTCTCTCTCTCTCTCT 418
Db 680 TGGGATATTTTCTTCTCTCTGAGAGCTGCTGTCGCTGCTCTCTCTCTCTCTCT 739
QY 419 GGGTGGAGGACCGGCTTCCGGAGTCTTCTGAGAGCAATCAAGTCTTACCGGAGATA 478
Db 740 GGATCAAGAGACGAGCTGATTTTCTTATTAACACACATCAGAGCTTACAGAGATGACA 799
QY 479 TCGATGCAAAACCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
Db 800 TTGATCTACAGAACTCATGAGCTTACCCAGGAAATTTGGAGTCTGAGGCTTTTG 859
QY 539 GCCCTGAAGCTGGGACCTCAACGCTACTTCAATTCAGCGGCTGCTGCTGCTGCT 598
Db 860 GAGCTGATGATGAGCTCAATATATTTACTTCAATTCAGCTGCTGCTGCTGCTGCT 919
QY 599 AGAAGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
Db 920 AGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
QY 659 CACAGTGTGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
Db 980 CTCAGTGTGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039
QY 719 CGAAGGCTGCTATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
Db 1040 CAAAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
QY 779 GCGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838
Db 1100 GTATTTTCATAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
QY 839 TCTAGACATCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
Db 1160 TGAAGTGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1188

RESULT 15
AF065389 1405 bp mRNA linear PRI 28-APR-2000
LOCUS Homo sapiens tetraspan NET-4 mRNA, complete cds.
DEFINITION AF065389
ACCESSION AF065389.1 GI:3152702
VERSION AF065389.1
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1405)
AUTHORS Serru,V., Dessen,P., Bouchelx,C. and Rubinstein,E.
TITLE Sequence and expression of seven new tetraspans
JOURNAL Blochm. Biophys. Acta 1478 (1), 159-163 (2000)
MEDLINE 20185353
PUBMED 10719184
REFERENCE 2 (bases 1 to 1405)
AUTHORS Rubinstein,E., Serru,V. and Bouchelx,C.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1998) INSERM U268, 14 av Paul Vaillant Couturier,
Villejuif 94807, France

FEATURES
source
1..1405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE ID 219547"
348..1154
/note="similar to Tspan-5; TM4SP"
/codon_start=1
/product="tetraspan NET-4"
/protein_id="AAC17120.1"
/db_xref="GI:3152703"
/translation="MSGKHYKGPVSCIKYIFGFENVIFWFLGITFLGILMANEK
GVLNLSITDLDGDPVWFLVGVGMFLIFGACIALRENTFLKFFSVFLGILF

FILETAGVIAEYFKDWIKDQLYEFINNIRAYRDDIDLONLIDFTOEYWCAGAGAD
DMNLIYFNCITDSNBSRESCVDPEDVINTQCGIDAROKPEVDQIVITIT
KCCVPOFEKWLQDNLIIYAGIFIGIALIIFIGICLAONLIVSDIEAVRASW"
BASE COUNT 309 a 357 c 387 g 352 t
ORIGIN

Query Match 24.3%; Score 337.4; DB 9; Length 1405;

Best Local Similarity 62.8%; Pred. No. 7.1e-67;

Matches 524; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

QY 33 GCGCGCGCGCGCGGATCTCTCTCTCGAGAGATGCATATTATAGATCTTAACGC 92
DB 314 GCTGCCGGGCAAGTGTCTTTCATGAACCCAGAGATGTCCGGAACTACAGAGGTCC 373
QY 93 CAAAGTCAGCTGCTGTGACAGTACCTCTTTTCAGCTACACATCATCTTGTGGC 152
DB 374 TGAAGTCAGTTGTGATCAATATCTTCAATTTGGCTTCAATGTCATATTTGGTTTTT 433
QY 153 TGGAGTGTCTCTCTTGGATCGGGCTGTGGGATGAGCGAAGAGGTGCTGCCGA 212
DB 434 GGGAAATTAATTTCTTGGAAATGGACTGTGGCATGGAATGAAGAGAGTCTGTCAA 493
QY 213 CCTCACCAGATGCCCGGATGCATGCAATCGACCCCTGTGTCTGTCTGTATGGTGG 272
DB 494 CATCTCTTCCATCACCATCTCGGGGCTTTGACCCAGTTTGGCTCTTCTGTGGTGG 553
QY 273 CGTGTGATGTTCCCTCGGGGTTCCCGGCTGTGGGCTGTGGGAGATATCTG 332
DB 554 AGGAGTATGTTCAATTTGGGATTTGCAGGGTGCATTTGAGCGCTACGGGAAAAACACTT 613
QY 333 CTTCCTCACTTTCTTCTGTGGCACATGTCATCTTCTTCTGTGAGCTGTGGTGGC 392
DB 614 CTTTCTCAAGTTTCTTCTGTGTCTGTGGCATTTATTTCTTCTGTGAGCTGTGGCGG 673
QY 393 CGTCTGTGCTTCTCTGTCCAGACTGGGTGAGGACCGGTTCCGGAGTTCTTCAGAG 452
DB 674 AGTTCTAGCATTTGTTCCTTCAAGAGCTGATCAAGACAGCTGATTTCTTTAATAACAA 733
QY 453 CAACATCAAGTCTACCGGAGCATGATGATCTGCAAAACCTCATGCACTCCCTTCGAA 512
DB 734 CAACATCAAGCATATCGGATGATGATTTGCAAAACCTCAWAGACTTCACCCAGGA 793
QY 513 AGCTAACAGTGTGTGGCATATGGCCCTGAGACTGGGACTCAACGTCCTACTCAA 572
DB 794 ATATTGGCAGTGTGTGGGCTTTTGGAGCTGATGATTTGAACTAATAATTTACTTCA 853
QY 573 TTGCAGCGGTCCAGCTACAGCCGAGAGAGTGGGGTCCCTTCTCTGTGTGTGCC 632
DB 854 TTGCACAGATTCATGCAAGTGCAGAGCATGTGGGCTTCTCATTTCTGTGTGACTAA 913
QY 633 AGATCTTGGCAAAAGTTTGAACACACAGTGGATGATGATGATGATGATGATGATGAA 692
DB 914 AGATCCCGAGAGATGTCATACACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 973
QY 693 GAGCAAGTGGATGATGATCTTTCACGAAAGGCTGATCCAGGCGCTGAAAGCTGGCT 752
DB 974 AGTTGACCAAGATTTGAATCTACAGAAAGGCTGTGTGCCAGTTTGAAGAGTGT 1033
QY 753 CCGCGGGAACATTTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
DB 1034 GCAGGACAAATTTAACTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1093
QY 813 TGGCATCTTCTTGGCAAGAGCATGATCAAGATGAGGAGTGAAGCGCGG 867
DB 1094 TGGGATATGCTGTGGCCAGAAATTTGTTAGCATATCGAAGCTGTCAAGGCGGAGC 1148

Search completed: December 19, 2002, 04:48:40
Job time : 2928 secs


```
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
US-08-408-222B-2

Query Match          5.0%; Score 69.8; DB 1; Length 687;
Best Local Similarity 53.1%; Pred. No. 1.8e-09;
Matches 199; Conservative 0; Mismatches 167; Indels 9; Gaps 2;
```

```
RESULT 5
US-08-408-222B-3
Sequence 3, Application US/08408222B
Patent No. 576727
GENERAL INFORMATION:
APPLICANT: Ikeyama, Shuichi
APPLICANT: Koyama, Masaru
APPLICANT: Miyake, Masayuki
APPLICANT: Senoo, Masaharu
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 112...795
OTHER INFORMATION: E Mat peptide
US-08-408-222B-3

Query Match          5.0%; Score 69.8; DB 1; Length 1120;
Best Local Similarity 53.1%; Pred. No. 2.2e-09;
Matches 199; Conservative 0; Mismatches 167; Indels 9; Gaps 2;
```



```

RESULT 8
US-08-254-493-3
; Sequence 3, Application US/08254493
; Patent No. 5439886
; GENERAL INFORMATION:
; APPLICANT: IKEYAMA, SHUICHI
; APPLICANT: KUYAMA, MASARU
; APPLICANT: MIYAKE, MASAYUKI
; APPLICANT: SENO, MASAHARU
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
; TITLE OF INVENTION: PRODUCTION THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/865552
; FILING DATE: 09-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 079996-1991
; FILING DATE: 12-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 085396-1991
; FILING DATE: 17-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 022321-1992
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 41777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; CELL TYPE: BREAST CARCINOMA
; CELL LINE: ZR-75-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..795
; FEATURE:
; NAME/KEY: mat -peptide
; LOCATION: 115..795
; US-08-254-493-3
;
; Query Match 4.7%; Score 65; DB 1; Length 1120;
; Best Local Similarity 52.3%; Pred. No. 4.3e-08;
; Matches 196; Conservative 0; Mismatches 170; Indels 9; Gaps 2;

```

```

QY 96 AGTCAGCTGCTGTACAGTACCTCTTTTCAGCTACACATCATCTTGTGGCTGG 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 AGGACACAGTGCATCAATATCTGCTGTTGGAATTAATCTTCTGCGCTGGCG 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 156 AGTTGCTTCCCTTGAGTGGGCTGTGGGCAATGAGGAAAGGCTGTCCAGCT 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 GATTGCTGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 216 CACCAAGTACCCGGATGATG-----GATGACCCCTGTGTGTGTGTGTGAT 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 CGACCAATAATATATATATATATATATATATATATATATATATATATAT 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 GGTGGCGGTGTGATGTTACCTGAGGATTCGCGGCTGCGGCGCTGCGGAGAA 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 AGCGGCGCCCTCATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 TATCTGCTGCTCAACTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 CCAGTGCATGCTGGGACTGTTCTTGGCTTCTTGTGTGTGTGTGTGTGTGTGT 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 387 TGTGGCCCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 446
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 TCGGCGCATGCGGATATTCACAGAGATGATGATGATGATGATGATGATGAT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 447 CGAGAGCAACATCAA 461
   ||| ||| ||| |||
Db 486 CAGGACACCTACAA 500
   ||| ||| ||| |||

```

```

RESULT 9
US-08-705-771-8
; Sequence 8, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MOLLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1344 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-705-771-8

```

2

For

104

27 C 3

515

552

575

5

612

RESULT 12
US-08-430-225A-19
Sequence 19, Application US/08430225A
Patent No. 6204000
GENERAL INFORMATION:
APPLICANT: Deng, Jin-Tang; Barrett,
APPLICANT: J. Carl; Lamb, Patricia W.; Isaacs, John T.
TITLE OF INVENTION: DIAGNOSTIC METHODS AND
TITLE OF INVENTION: GENE THERAPY USING REAGENTS DERIVED FROM THE
TITLE OF INVENTION: HUMAN METASTASIS SUPPRESSOR GENE KAI1
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

RESULT 13
US-08-807-044-4
Sequence 4, Application US/08807044
Patent No. 5863735
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K
TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/807,044
FILING DATE: Herewith
CLASSIFICATION: 530

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: MakI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 801 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-020-956-16

Query Match	Similarity	66.9%	Score 362	DB 9	Length 813
Best Local	Similarity	66.9%	Pred. No. 3,6e-88		
Matches 315	Conservative	0	Mismatches 255	Indels 0	Gaps
QY	55	AGATGAGTGCCTGCTGATACAGTACCTCCCTTTTACGCTCAACATCATCTCTGGTGGCTG	154		
Db	35	AGGTGCGAGTGCCTGCGGGAATACCTTCCTGTTGACCTTCAACATGTCTCTGGTGTCTGG	94		
QY	155	GAGTGTCTTCTTGAGTTCGGGCTGTGGCATGAGACGAAAGGTGTGCTGTCCGACC	214		
Db	95	GAACTGTCTTCTTGAGTTCGGGCTGTGGCATGAGACGAAAGGTGTGCTGTCTCCGACC	214		
QY	215	TCACCAAGTGCACCGGATGATGGAATTCGACCTGTGGTGGTGGCTCGATGATGGGCG	274		
Db	155	TCACCAAGTGCACCGGATGATGGAATTCGACCTGTGGTGGTGGCTCGATGATGGGCG	274		
QY	275	TGATGATGTTCACCTCGGGGTTCCCGGCTGCTGGGGGCTCGCGGGAGAAATATCTGCT	334		
Db	215	GCTCTCATGTGCGGTGTGGGCTTTTCTGTGGCTGCAATTTGGGGCCCTCCGGGAGAAACCTTCC	274		
QY	335	TGCTCAACTTTTCTGTGGACCAATCTGCTCATCTTCTCCGGAGGCGGGGTGGGCG	394		

Db	275	TGCTCAAGTCTTTCCTCCGTCTCTCTCGGCTCACTCTTCTCTGAGACTGGCAACAGGCA	334
Qy	395	TGCTGAGCCTTCTCTTCCAGAGCTGGGTGAGAGACCGGTTCCGGAGATCTTGTGAGAGCA	454
Db	335	TCTCTGGCCTTGTCTTCAAGSACCTGGANTTGGAGACCAGCTCAACCTCTTATCAACAGACA	394
Qy	455	ACATCAAGTCTCTACCGGAGCATATGCATCTGCAAAACCTCATCTACTCCCTTCAGAAAG	514
Db	395	ACGTCAAGGCTTACCGGGAGCAGCATTTGACCTCCAGAACCTCATTTGACTTGGCTAGGAAT	454
Qy	515	CTAACAGTGTGTGGGGGCATATGGCCCTGAGSACTGGAGCCTCAAGCTCATCTCAATT	574
Db	455	ACTGCTCTTGTGTGGAGCCCGAGGGCCCAATGCTGGAACCTCAATATCTACTTCACT	514
Qy	575	GCAGCGGTGCCAGCTACAGCCGAGAGAAAGTGGCGGGGTCCCTTCTCTGCTGGCTGCAG	634
Db	515	GCACGACTTAAACCCAGCCGGGAGGGGTGGGGGGTCCCTTCTCTGCTGGCTGCAGG	574
Qy	635	ATCTCTGGCAAAAAGTTGTGAAACAACACAGTGTGGATATGATGTCAGATTCAGCTTAAGA	694
Db	575	ACCTCTGGAGAGATGCTCTTCAACACCCAGATGTGGCTACGAGTCCGGCTCAAACTGGAGC	634
Qy	695	GCAGTGGGATGATGCCATCTTTCAGSAAAGGCTGCATTCACAGGGCTGGAAAGCTGGCTCC	754
Db	635	TGGACACACAGGGCTTCACTCACACCAAAGGCTCGTGGGCCAATTTGAGAGAATGGCTGC	694
Qy	755	CGCGGAACATTTACATTGTGGCTGGCTGGCTTCATCGCCATCTCCCTGTGCGAATATTG	814
Db	695	AGGACAAACCTGATTTGTGGTGGGGGAGTCTTCAATGGGAGATGGCCCTCCTCCAGATCTTTG	754
Qy	815	GCATCTTCTTGGCAAGACGCTGATCTTCAGACATTCGAGGCAAGTAAAGGC	864
Db	755	GCATCTGCTTGGCCAGAACCTCGTGAATGATCATCAAGGCAAGTAAAGGC	804

```

RESULT 2
US-09-934-268-1
; Sequence 1, Application US/09934268
; Patent No. US20020172986A1
; GENERAL INFORMATION:
; APPLICANT: Leidy, Kevin R.
; TITLE OF INVENTION: 23228, A NOVEL HUMAN TETRASPANIN FAMILY
; TITLE OF INVENTION: MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-079001
; CURRENT APPLICATION NUMBER: US/09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3184
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (168)...(977)
US-09-934-268-1

Query Match      26.1% Score 362; DB 9; Length 3184;
Best Local Similarity 66.9%; Pred. NO. 6.8e-88;
Matches 515; Conservative 0; Mismatches 255; Indels 0; Gaps

QY 95 AAGTAGCTGCTGTACAGTAGTACCTCTTTTCAGCTACAACATCATCTTCTGTGGCTG 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 202 AGGTGGGCTGCTGGCGGGAATACTTCTGTTGGCTTCAACATGTCTTCTGGGTCGTG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 155 GAGTGTCTTCTCTTGAGATCCGGCTGTGGCGCATGGAGCGAAAGGTGTCTCTCCAGC 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 262 GAGCCCTCTTCTGGCTATGAGCCTCTGGGCTTGGGGTGAAGAGGGCGTTCTTCGACAC 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 215 TCACCAAGTACCCGGATGATGGAATCGAACCCCTGTGCTCTGCTGATGATGGTGGCG 274

```

Db	322	TCCTACGCCCTCAGCAGATCTGGAGAGCCTTACCCCGTGTGGCTTTGTGTGATGTTGGAG	381
Qy	275	TGCTGATGTTTCACCCCTGGGGTTGCGCGGCTTCGTGGGGCTCTCCGGGAGATATCTGCT	334
Db	382	GCCTCATGTCTCGTCTGGGGCTTTGCTCGTCGATTTGGGGCCCTCCGGGAAACACCTTCC	441
Qy	335	TGCTCACTTTTCTGTGGGACCAATCTGTCATCTTCTTCGAGAGCTGGCTGTGGCG	394
Db	442	TGCTCAAGTTTTTCTCCGTCTCTCGGTCTCATCTTCTTCTCGAGACTGGCAACAGGA	501
Qy	395	TGCTGGCCCTTCCTGTCCAGGACTGGGTGAGGGACCGGTTCCGGAGTCTTCTGAGACA	454
Db	502	TCTGGCCTTGTCTTCAAGGACGTGATTTGAGACACGCTCAACCTTCTATCAACAACA	561
Qy	455	ACATCAAGTCTTACCGGGAGATTTGATCTGCAGAAACCTCATCGACTCCCTTCAGAAAG	514
Db	562	ACGTCAAGGCTTACCGGGAGCACAATTAACTCCAGAACCTCATTTGACTTTGCTCAGGAAT	621
Qy	515	CTAACCACTGCTGTGGGCGATATGTGGCCCTGAAGACTGGGACCTCAACGTCYACTTAATT	574
Db	622	ACTGGTCTTGGTGTGGAGCCCGAGGGCCCAATGACTGGAACCTCAATATCTACTTCAACT	681
Qy	575	GCAGCGGTGCCAGCTACAGCCGGAAGAAGTCCGGGGTCCCTTCTCTGTGGTGGCAG	634
Db	682	GCACCTGACTTGAACCCGACCGGAGAGGCTTCGGGGGTGCCCTTCTCTGTGGTCAAGG	741
Qy	635	ATCCTGGCCAAAGATTGTGAMCACACAGTGTGATTTGATGTCAAGATTCAAGCTGAGA	694
Db	742	AACCTGGCGGAGGATGTCTCAACACCCAGTGTGGCTAGAGCTCCGGCTCAAACTGGAGC	801
Qy	695	GCAMTGGGATGAGTCCATCTTTCACGAAAGGCTCTCATTCAGGGGCTGGAAAGCTGGCTC	754
Db	802	TGGAGCACACAGGGCTTATTCACACACAAAGGCTCGTGGGCCAAGTTTGAAGAGTGGCTC	861
Qy	755	CGCGGAACATTTAATTGTGCTGGCGCTTTCATTCGCATCTCGCTTTCAGATATTTT	814
Db	862	AGGACAACCTGATTTGTGTGGCGGAGATCTTCATTTGGGCACTGGCCCTCTCCAGATCTTTG	921
Qy	815	GCATCTTCTTGGCAGAGCAGCTGATCTTCAGACATTCAGAGCAGTGAAGCC	864
Db	922	GCATCTGCTTGGCCCGAGAACCTCTGTAGGTGACATTCAGGACAGTGTAAAGCC	971

```

RESULT 3
US-09-972-970-2
; Sequence 2, Application US/09972970
; Patent No. US2002016493A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: TM4SF Receptor Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P050691
; CURRENT APPLICATION NUMBER: US/09/972,970
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11130
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,336
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-970-2

Query Match      25.3%; Score 351.8; DB 9; Length 2538;
Best Local Similarity 66.9%; Pred. No. 3,4e-85;
Matches 517; Conservative 0; Mismatches 252; Indels 4; Gaps 1;

QY      92  CCAAGTCAGCTGCGGTCGACAGTAGTACCTCCTTTCACTCACTCAACATCATCATCTGCGTTGG 151
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      150  CCGAGGTGGCGGTCGCTGCGGGAATACCTCCTGTTGGCTTCACACATATGTTTCTGGGTGC 209

```

```

OY 152 CTGAGTGTCTCTCTTGGAGTCGGGCTGTGGCATGAGCAAAAGGGTGTCTCG 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 TGGAGACCTGTCTCGGCATCGGCTCTGTGGCTTGGAGGTGAGAGGGTGTCTCCA 269
OY 212 ACCTACCAAAAGTACCCCGATGATGAATGACCTTGTGTGTCTGATGAGTGG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ACATCTCTCGCTGACGATCTGGGAGGCTGACCCCTGTGGCTGTGAGTGGTGG 329
OY 272 GCGTGTGATGTTCACCTTGGGGTGTGGCGCTGGCTGGGCTGTGGCGAGAAATCT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 GAGGCTGTATGTCTGCTGTGGGCTTGGCGGCTGTGAGGCTGTCCGGGAGAACAT 389
OY 332 GCTTGTCAACTTTTCTGTGGCACCATGCTGTCTCTCTCTGAGTGTGGCTGG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 TCTGTCTCAAGTTTCTGATGTCTTGGCTTGGCTTCTCTCTGAGTGTGGCAACG 449
OY 392 CCGTGTGGCTTCTGTTCACGAGTGTGGTGGAGGAGGCTTCCGGAGTCTTTCGAG 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 GGAATCTGGCTTCTGTTCACGAGTGTGGTGGAGGAGGCTTCAATTTCTTCAATTA 509
OY 452 GCAACATCAAGTCTACCGGAGCATATGATCTGCAGAAACCTCATCGACTCCCTTCA 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ACAACGTCAAGGCTATCGGAGATGACATGACCTCAAGAACCTTGAATTTGCTCAG 569
OY 512 AAGCTAACAGTGTGTGGCGCATATGCCCCCTGAAGACTGTGGACCTCAAGTCTACT 571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 AATATGTCTTGTCTGTGGAGCCCGAGGCGCTTAAATGACTGGAACCTCAATATCTAT 629
OY 572 ATTGCAACGGGTGCGAGCTACAGCGAGAGAGAGGCGGGGCTCCCTCTCTCTGTGGTGC 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 ACTGCACTGACTTGAACCCAGCCAGAGAGCGCTGCGGGGTGCCCTTCTCTGTGTGTA 689
OY 632 CAGATCTGTGGCAAAAAGTTGTGACACACAGTGTGGATGATGTGAGATTTCACTGA 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 GGGACCTGGG---ATGTCTCAACACCCAGAGTGTGGCTATGATGTGCGGCTCAAACTGG 745
OY 692 AGACCAAGTGGAGATCTCATCTTCAAGAAAGCTGCATCCAGGGCTGTGAAGCTGGC 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 AGCTGAGACAGAGGCTCTATACACCAAGAGCTGTGTGGCCATTTGAGAAAGTGGC 805
OY 752 TTCCGCGAACATTTACATTTGTGGCTGGCTCTTCATCGCATCTGCTTTCAGATAT 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 TGCAGGACACATGATGTGTGTGGCTGGGCTTTGTGGGCAATGCTCTCTCAAGATCT 865
OY 812 TTGGCATCTTCTGTGGCAAGACGCTGATCTCAGACATCGAGCGAGTGAAGGCC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 TTGGATCTGTGCTGGCCAGAACCTTGTGATGATCAAGGAGTGAAGGCC 918

```

RESULT 4

```

US-09-925-299-205
: Sequence 205, Application US/09925299
: Patent No. US20020055627A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102
: CURRENT APPLICATION NUMBER: US/09/925, 299
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124, 270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1556
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 205
: LENGTH: 1655
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1548)

```

```

: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1559)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1564)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1623)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (1643)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-205

Query Match      24.38; Score 337.4; DB 10; Length 1655;
Best Local Similarity 62.8%; Pred. No. 2,2e-81;
Matches 524; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

OY 33 GCGCGCGCGCGCGGAGTCTGCTTCTCAGAGATGCACTATTAATGATCTTAACGC 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GCTGCGGGGCAAGTGTCTTATGATGACCAAGATGTCCGGAGAGCAAGGATGCC 302
OY 93 CAAAGTCAAGTCTGTGTGACAAAGTACCTCTTTCAGTACAAATCATCTTGTGGTGGC 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 TGAAGTCAAGTGTGTGATCAATCAATTAATCTTCAATTTGGCTTCAATATTTGGTTT 362
OY 153 TGAAGTGTCTTCTGTGTGAGTGTGGGCTGTGGGCAATGAGCAAAAGGCTGTGCGA 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 GGAATTAACATTTTGTGAATGTGACTGTGGCATGAGATGAAGAGAGTTGTGTCCA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 213 CCTCAACAAAGTACCGGATGATGATGATGACCTGTGTGTGTCTGTATGTGTGG 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 CATCTCTTCAATCACCAGATCTGTGGGCTTGTACCCATTTGGCTCTCTGTGTGG 482
OY 273 CGTGTGATGTTCACCTGTGGGCTGTGCGGCTGTGGGCTGTGGGAGATATGTG 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 AGAGAGTATGATTTGATTTGGATTTGACAGGCTGTGATGAGGCTACAGGAAACCTTT 542
OY 333 CTTCGCACTTTTCTGTGTGGCACCATGTGTCTCATCTTCTGTGAGTGTGGC 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 CCTTCAGATTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
OY 393 CGTGTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 AGTTTACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 662
OY 453 CAACATCAAGTCTTACCGGAGATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 CAACATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
OY 513 AGCTAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 ATATGTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 782
OY 573 TTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 TTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
OY 633 AGATCTGTGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 AGATCTGTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
OY 693 GAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 AGTTGACAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
OY 753 CCGCGGACATTTTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GCAAGCAATTTTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
OY 813 TTGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db 374 TGAAGTCAGTGTGGATCAAAATACATATTGGCTTCATATGTCATATTGTTT 433
QY 153 TGAAGTGTCTTCCTGGAGTCGGCTGTGGCATGAGAGAAAAGGTGTCTGCCA 212
Db 434 GGGAAATACATTTCTGGAAATGAGCTGGCATGGAATGAAGAGGTTCTGCCAA 493
QY 213 CCTCAACAAAGTGAACCCGGATGATGGAATGACCCCTGTGTCTGCTCATAGTGG 272
Db 494 CATCTCTTCATCAACCCGATCTCGGGCTTTGACCCAGTTTGGCTCTCTCTGGTGG 553
QY 273 CGTGTGATGTTCAACCCGGGTTCCGGCTGGCGGGGCTCTCGGGAGAAATCTG 332
Db 554 AGGAGTGTGTTCAATTTGGGATTGGAGGTCATTTGAGCGCTCAAGGAAACCTTT 613
QY 333 CTTCCTCAACTTTCTGTGGCACCACATCTGTCATCTTCTCTGGAGCTGTGGC 392
Db 614 CCTCTCAAGTTTCTGTGTCTCGGGAAATTTTCTCTGGAGCTCACTGCCG 673
QY 393 CGTGTGGCTTCTGTGTCCAGAGCTGGTGAGGACCGGTTCCGGGAGTTCTTGAG 452
Db 674 ACTTCTACATTTGTTTCAAGAGCTGATCAAGACACCTGTATTTCTTATAACAA 733
QY 453 CAACATCAAGTCTCACCGGACATGATGATCAACAAACCTCATGACTCCCTCAGA 512
Db 734 CAACATCAGAGCATATGCGGATGATGATGATTTGCAAAAACCTCATGACTCAG 793
QY 513 ACCTAACAGTGTGTGGCGCATATGCGCCCTGAAGAGCTGAGACCTCAACGCTCT 572
Db 794 AATTTGAGGATGTGTGGGCTTTTGGAGCTGATGATGGAACCTCAAAATATTACT 853
QY 573 TTGCAGCGGT 582
Db 854 TTGCACAGAT 863

```

```

RESULT 7
US-09-729-674-169
; Sequence 169, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO: 169
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-169

```

```

Query Match 16.5%; Score 229; DB 10; Length 1110;
Best Local Similarity 57.1%; Pred. No. 3.1e-52;
Matches 477; Conservative 2; Mismatches 292; Indels 64; Gaps 1;
QY 33 GCCGCCGCCGCCGCCGCTTCTGCTTCTCAGAGATGCACTAATATAGATACTCTAACGC 92

```

```

Db 142 GCTGCCGGGGGCAAGTGTCTTCTATGAAACCCAGAGATGTCCGGGAAAGCACTACAGGGTCC 201
QY 93 CAAAGTCACTGCTGTACAGTACCTCTTTTACGTACACATCATCTTCTGTGGTGC 152
Db 202 TGAAGTCAGTGTGTATCAATTAATCTTATTTGGCTTCAATATTTGTTGGTTT 261
QY 153 TGAAGTGTCTTCCTGGAGTGGGCTGGGCTGGAGAGGAAAGAGTGTGTGCCA 212
Db 262 GGGAAATACATTTCTTGAATTTGGACTGTGGCATGGAATGAAGAAAGATTTCTGCCA 321
QY 213 CCTCAACAAAGTGAACCCGGATGATGGAATGACCCCTGTGTCTGCTCATAGTGG 272
Db 322 CATCTCTTCATCAACCCGATCTCGGGCTTTGACCCAGTTTGGCTCTCTTGTGGTGG 381
QY 273 CGTGTGATGTTCAACCCGGGTTCCGGGCTGGGCTGTGGGAGATATCTG 332
Db 382 AGGAGTGTGTTCAATTTGGGATTTGGATTTGCAAGGTTGATGAGCGCTACGGGAAACACTT 441
QY 333 CTTCCTCAACTTTTCTGTGGCACCATGCTCATCTTCTCTGAGCTGTGCTGCC 392
Db 442 CTTCCTCAAGTTTCTGTGTCTGTGGAAATTTTCTCTCGGAGCTCACTGCCG 501
QY 393 CGTGTGGCTTCTGTGTCCAGAGCTGGGTGAGGAGGACCGGTCCGGAGTTCTTGAG 452
Db 502 AGTTCTAGCATTTGTTTCAAGAGCTGATCAAGACCACTGATTTCTTTATAACAA 561
QY 453 CAACATCAAGTCTCACCGGAGATATGATCTGCMAAACCTCATGACTCCCTCAGA 512
Db 562 CAACATCAAGCATATGCGGATGATGATGATTTTCAAAACCTCATGACTCAGCA 621
QY 513 AGCTAACAGTGTGTGGGATGATGCGCTGAGAGTGGGAGCTCAACGCTACTTCAA 572
Db 622 ATATA----- 626
QY 573 TTGCAGCGGTGCCAGCTACAGCCGAGAGAGTGGGGTCCCTCTCTCTGCTGCC 632
Db 627 -----TCCAAATGCAAGTGCAGAGCATGTGGGCTTCATTTCTCTGTGCATA 677
QY 633 AGATCCTGGGCAAAAAGTGTGACACACAGTGGGATGATGATGATGATGATGATG 692
Db 678 AGATCCCGAGAGATGTATACACATGAGTGGCTATGATGATGATGATGATGATG 737
QY 693 GAGCAAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 752
Db 738 AGTTGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 797
QY 753 CCGCGGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 812
Db 798 GCAGGACAAATTTACCCWGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 857
QY 813 TGGCATCTCTCTGGGCAAGGACGCTGATCTCAGACATCAGAGCAGTGAAGCGCGC 867
Db 858 TGGGATMTGCTGGGCCCAAGATTTGTTAGGATATGATGATGATGATGATGATG 912

```

```

RESULT 8
US-09-764-873-12
; Sequence 12, Application US/09764873
; Patent No. US20020151479A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P201
; CURRENT APPLICATION NUMBER: US/09/764,873
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO: 12
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-764-873-12

Query Match 11.9%: Score 165.2; DB 10; Length 545;
Best Local Similarity 98.2%; Pred. No. 3.6e-35;
Matches 167; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CTCGGCCGACCGCGCGCGCGCGCTGCCGCCGCCGCCGGGATTCCTCTCAGA 64
Dy 9 CTCGGCCGACCGCGCGCGCGCGCTGCCGCCGCCGCCGGGATTCCTCTCAGA 68
Qy 65 AGATGCACTATTAGTACTCTAAGCAAGTACAGCTGCTGACAGTACCTCTT 124
Dy 69 AGATGCACTATTAGTACTCTAAGCAAGTACAGCTGCTGACAGTACCTCTT 128
Qy 125 TCACCTACACATCATCTTCTGTTGGCTGAGTTGTTCTCTTGAGTCTC 174
Dy 129 TCACCTACACATCATCTTCTGTTGGCTGAGTTGTTCTCTTGAGTCTC 178

RESULT 9

US-10-052-586-277
Sequence 277, Application US/10052586
Patent No. US20020127584A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C1

CURRENT APPLICATION NUMBER: US/10/052.586
CURRENT FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063490
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772

PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/07450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15

; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: 60/089908
 Query Match 8.8%; Score 122.8; DB 12; Length 1778;
 Best Local Similarity 52.5%; Pred. No. 1,7e-23;
 Matches 321; Conservative 0; Mismatches 282; Indels 9; Gaps 2;

```

QY 242 TCACACCTGTGTGCTGTGCTCTGATGTGGGCGTGTGTGATGTTCACCTGGGCTGGCG 301
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 373 TGGCTCCAGCCCATATCTCTATCTCTCCCTGGGCGTGTGTGATGTTCATGTCTCCATFTG 432
QY 302 GCTGTGTGGGGGCTGTGGGGGAGAAFTMTGCTGTGCTCACTTTTCTGTGGACCATGCG 361
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 433 GTGTGTGTGCGTCTCCCTCCGTGTGACACTGTGACTTCTTCACAGCATTTATGTACTTCCTTG 492
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 362 TGTCTATCTTCTTCTGAGCTGTGGCGTGTGGCGGCTGTGGCGCTTCTCTGTTCACGAGCTGGG 421
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 493 GGATCTGCTCATGTATGTAGAGCTCATGTGTGGCGGTGTGGCGCTTGACCTTCGGMACCAGA 552
QY 422 TGAGAGACCGGTGTCCGGGAGTTCCTTCAGAGACAACTCAAGTCTTACCGGGAGCATATCG 481
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 553 CCATTGACTTCTCTTAACGACAACTTGTGAAGAGGAAATGTGAACCTATGATGATCTGTG 612
QY 482 ATTCGCAAAACCTCATGCACTCCCTTCAGAAAAGCTAACGAGTGTGGGCGCATATGGCC 541
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 613 ACTTCAAAACATCATGTAGACTTTTGTGAGAAAAGTTCMAGTGTGTGGCGGGAGAGACT 672
QY 542 CTGAAGACTGTGGACCTCAACGCTTACTTCAATTGTGCAGCGGTGCCAGCTAACGCCAGAGA 601
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 673 ACCGAGATTGTGAGCAAGAAATCATGATCCACGACACTGCACTG-----CCCTGTGACCCCTGG 726
QY 602 AGTGCAGGGGTCCCTTCTCCGTGCGTGCAGATCTCTGGCAAAAAGTTGTGAACACAC 661
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 727 CCTGTGGGTGCTCCTTAACACTGCTGTGACTCAGAA---CAGCAGACAGAGTTGTCAACACCA 783
QY 662 AGTGTGATATGATGTGAGGATTCAGCTGTGAAGACAGAGTGGGATGTAGTCCATCTTACGA 721
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 784 TGTGTGCTACAAAACATATGACAAAGAGCGTTTCACTGTGACGAGATGATCATCTAGCTGC 843
QY 722 AAGGCTGCATTCAGGCGCCTGGAAAGCTGGCTCCCGCGGAACATTTACATTTGTGGCTGGC 781
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 844 GGGGCTGCACCAACGCGGTGATCATCTGTGTTATGTAGCAACTCAACCATCATGTGGCGTGA 903
QY 782 TCTTCATGCGCATCTCGCTGTGAGATATTTGGCATCTTCTGGCAAGGACGCTGATCT 841
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 904 TCTCTCGGGGATCTGTGCTTCCCAAGTTCTCGGGGGGTGTGTGAGCGTGTCTTACATCA 963
QY 842 CAGACATCGAGG 853
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 964 CCCGGGTGAGG 975
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-10-042-417-37
; Sequence 37, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
  
```


QY 447 CGAGAGCAACATCAA 461
| | | | |
Db 475 CAAGGACACCTTACAA 489

Search completed: December 19, 2002, 06:22:12
Job time : 177 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 19, 2002, 05:20:37 ; Search time 2589 seconds

(without alignments)
3035.055 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYRNAKSCWYKYLFL.....IFLARLISDIENAKAGHHF 270

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO/US09905674/runat_16122002_111845_6756/app.query.fasta_1.455
-DB=GenEmbl -QFMT=fastlap -SUFFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCCALIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEPSTLZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905674.ecgn_1_1.1616-ernat_16122002_111845_6756 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pin:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	100.0	1388	6 AX440923	AX440923 Sequence
2	1449	100.0	2553	9 AF311903	AF311903 Homo sapi
3	1438	99.2	2556	9 HSM801611	AL136638 Homo sapi
4	1420	98.0	2428	10 BC025568	BC025568 Mus muscu
5	1420	98.0	2498	10 BC024611	BC024611 Mus muscu
6	1420	98.0	2500	10 BC026574	BC026574 Mus muscu
7	1338.5	92.4	2502	9 BC002920	BC002920 Homo sapi
8	867	59.8	1405	9 AF065389	AF065389 Homo sapi
9	867	59.8	1416	9 BC009704	BC009704 Homo sapi
10	867	59.8	3175	10 AF121344	AF121344 Mus muscu
11	861	59.4	1516	10 BC010346	BC010346 Mus muscu
12	858	59.2	813	6 AX420468	AX420468 Sequence
13	858	59.2	3184	6 AX420466	AX420466 Sequence
14	849	58.6	1408	9 AF053455	AF053455 Homo sapi
15	825	56.9	2426	9 BC010405	BC010405 Homo sapi
16	572	39.5	1995	10 BC024685	BC024685 Mus muscu
17	568	39.2	852	6 AX247836	AX247836 Sequence
18	567	39.1	864	6 AX343015	AX343015 Sequence
19	559	38.6	1567	6 AX061778	AX061778 Sequence
20	522.5	36.1	102468	2 AC009984	AC009984 Drosophila
21	522.5	36.1	106584	2 AC013977	AC013977 Drosophila
22	522.5	36.1	140973	3 AC008140	AC008140 Drosophila
23	522.5	36.1	159007	3 AC009219	AC009219 Drosophila
24	522.5	36.1	224795	3 AE003688	AE003688 Drosophila
25	515	35.5	4445	9 AK024427	AK024427 Homo sapi
26	508	35.1	140468	2 AC017377	AC017377 Homo sapi
27	508	35.1	174311	3 AC0092216	AC0092216 Drosophila
28	508	35.1	174832	3 AC004758	AC004758 Drosophila
29	508	35.1	260550	3 AE003612	AE003612 Drosophila
30	503	34.7	1216	3 AF274013	AF274013 Drosophila
31	451.5	31.2	22920	3 CERI4G10	Z68880 Caenorhabditis
32	433	29.9	11172	2 AC013030	AC013030 Drosophila
33	433	29.9	191342	3 AC104140	AC104140 Drosophila
34	433	29.9	289090	3 AE003424	AE003424 Drosophila
35	406	28.0	56820	3 DMBR7C10	AF121804 Drosophila
36	404.5	27.9	1703	9 AF120266	AF120266 Homo sapi
37	404.5	27.9	1726	9 BC003157	BC003157 Homo sapi
38	404.5	27.9	1734	9 BC004161	BC004161 Homo sapi
39	395.5	27.3	1778	6 AX376210	AX376210 Sequence
40	393	27.1	1661	10 BC003872	BC003872 Mus muscu
41	385	26.6	1646	10 BC019991	BC019991 Mus muscu
42	370	25.5	1601	9 AF325213	AF325213 Homo sapi
43	370	25.5	1693	9 BC032802	BC032802 Homo sapi
44	349	24.1	557	6 AX079573	AX079573 Sequence
45	343.5	23.7	860	10 RNAD1A	X61654 Rat mRNA fo

RESULT 1

ALIGNMENTS

AX440923 1388 bp DNA linear PAT 28-JUN-2002
LOCUS AX440923
DEFINITION Sequence 1 from Patent WO0206340.
ACCESSION AX440923
VERSION AX440923.1 GI:21665563
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Reinhard, C.J. and Garcia, P.D.
Tetraspan protein and uses thereof
Patent: WO 0206340-A 1 24-JAN-2002;
CHIRON CORPORATION (US)
FEATURES
source 1.1388
Location/Qualifiers
BASE COUNT 270 a 390 c 411 g 313 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 6,37e-154 Length: 1388
Score: 1449.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-905-674-2 (1-270) x AX440923 (1-1388)
QY 1 MethISTYTYRATGTYRSerASnaAlaLysValSerCysTRPtyrLysTYrLeuPhe 20
|||||
Db 67 ATGCACATATATAGATACCTTAACGCCAAGTCAGCTGGTACAAAGTACCTCTTTC 126
|||||
QY 21 SerTYrAsnIlellePheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAla 40
|||||
Db 127 AGCTACAAACATCATCTTCTGGTGGCTGGAGTTGTTCTCTGGAGTCCGGCTGGGGCA 186
|||||
QY 41 TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
|||||
Db 187 TGGAGCGAAAGGCTGTGCTGCTCCGACCTCACCAGTACCCGGATGATGATCGAC 246
|||||
QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
|||||
Db 247 CCTGTGCTGCTGCTGCTGATGAGTGGCGGCTGGATGTCACCTGGCGGCTTC 306
|||||
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
|||||
Db 307 GTGGGGGCTGCGGAGAAATATCTCTTCTCACTTTCTGTGGCCATCTGCTGC 366
|||||
QY 101 IlePhePheLeuGluLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120
|||||
Db 367 ATCTTCTTCTGAGCTGGGCTGTGGCGCTGGCTTCTCTGTCAGAGACGCGGTGAGG 426
|||||
QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYrArgAspIleAspLeu 140
|||||
Db 427 GACCGGTTCCGGGAGTTCTTCGAGCAACATCAAGTCTCCTACCGGAGCATATCATCTG 486
|||||
QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyValAlaTYrGlyProGlu 160
|||||
Db 487 CAATACCTATCGACTCCCTTCAGAAAGCTAACCAAGTGTGGGCGCATATGCCCTGAA 546
|||||
QY 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYrSerArgLysCys 180
|||||
Db 547 GACTGGAGCTCAAGCTACTTCAATTCAGCGGTGCCAGTACAGCCGAGAGAGAGTGC 606
|||||
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200
|||||
Db 607 GGGGTCCTCTTCTCCGTGCGGCGCAGATCCTGCGCAAAAGTGTGACACACAGTGT 666
|||||
QY 201 GlyTYrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
|||||

|||||
Db 667 GGATATGATGTCAGATTACAGTGAAGGACAGTGGATGATGTCATCTTCACGAAAGGC 726
|||||
QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240
|||||
Db 727 TGCATCCAGGGCGCTGGAAGCTGGCTCCCGGGAACATTCATTTGTGGCTGGCTTTC 786
|||||
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
|||||
Db 787 ATCGGCATCTCGCTTGTGACATATTTGGCATCTTCTCTGGCAGAGACGCTATCTCGAC 846
|||||
QY 261 IleGluAlaValLysAlaGlyHisHisPhe 270
|||||
Db 847 ATCGAGCGAGTGAAGCCGCGCATCATCTTC 876
|||||
RESULT 2
AF311903
LOCUS AF311903 2553 bp mRNA linear PRI 16-AUG-2002
DEFINITION Homo sapiens DC-TM4F2 precursor, mRNA, complete cds.
ACCESSION AF311903
VERSION AF311903.1 GI:22266721
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 2553)
AUTHORS Zhang, W., Li, N., Man, T., and Cao, X.
TITLE Identification of novel membrane proteins
JOURNAL Unpublished
2 (bases 1 to 2553)
REFERENCE
AUTHORS Zhang, W., Li, N., Man, T., and Cao, X.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-2000) Department of Immunology, Second Military Medical University & Shanghai Brilliance Biotechnology Institute, 800 Xiangyin Rd., Shanghai 200453, P.R.China
FEATURES
source 1.2553
Location/Qualifiers
CDS
1.2553
/organism="Homo sapiens"
/db_xref="taxon:9606"
73.885
/codon_start=1
/product="DC-TM4F2 precursor"
/protein_id="AA094899.1"
/translation="MHYRYSNAKVSQWRYLLFSYNIIFMLAGVFLGLWAMSEK
GVLSDLTQVYRMGIDPVYVILMVGVMITIGAGCGVGLAREICLINFPGTYILIF
FLRLAVALVAFLEFDQWYRDRFRFFESNFKSYRDDIDLDNLDSKANCCGATGPE
DMDLNYFNCSGASYSREKGVFSCVDPAPKVVNTCCGIDVRIQLSKWDESLFT
KGCIGALESVLPNIIYIVAGVFIAISLDIIFGIFLARTLISDIEAVKQHP"
BASE COUNT 505 a 724 c 680 g 644 t
ORIGIN
Alignment Scores:
Pred. NO.: 1.4e-153 Length: 2553
Score: 1449.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-905-674-2 (1-270) x AF311903 (1-2553)
QY 1 MethISTYTYRATGTYRSerASnaAlaLysValSerCysTRPtyrLysTYrLeuPhe 20
|||||
Db 73 ATGCACATATATAGATACCTTAACGCCAAGTACGCTGGTACAAAGTACCTCTTTC 132
|||||
QY 21 SerTYrAsnIlellePheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAla 40
|||||
Db 133 AGCTACAAACATCATCTTCTGGTGGCTGGAGTTGTTCTCTGGAGTCCGGCTGGGGCA 192
|||||
QY 41 TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
|||||


```

Db      193 TGGAGCGAAAAGGTGTGCTGTCCGACCTCACCAAAAGTGACCCGAGTCATGGAATCGAC 252
Qy      61 ProValValIleuValIleuMetValGlyValIleuPheThrLeuGlyPheAlaGlyCys 80
Db      253 CCTGTGTCTGCTGCTGTCTGTATGTGGCGGTGTGTATGTTTCACCCCTGTGGGTTTCCCGGGCTGC 312
Qy      81 ValGlyAlaLeuAurGluAsnIleCysLeuLeuAsnPhenPheCysGlyThrIleValLeu 100
Db      313 GTGGGGGCTCTCGGGAGAAATATGTGCTTCAACTTTTCTGTGGCACCATGCTGCTC 372
Qy      101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArg 120
Db      373 ATCTTCTTCTCTGAGCTGTGCTGTGGCCGTGCTTCTGTTCTGACAGACTGGGTGAGG 432
Qy      121 AspArgPheArgGluPhePheGluSerAsnIleCysSerTyrArgAspAspIleAspLeu 140
Db      433 GACCGGTTCCGGAGTTCTTCAGAGCAACATCAAGTCTTACCCGGAGCATATCGATCTG 492
Qy      141 GlnAsnLeuIleAspSerLeuGlnIleLysAlaAsnGlnCysGlyAlaIleTyrGlyProGlu 160
Db      493 CAAACCTCATCGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGGCCCTGAA 552
Qy      161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
Db      553 GACTGGGACCTCAACGCTCTACTTCAATTGCACGGGTGCAGCTACAGCCGAGAGAAAGTGC 612
Qy      181 GlyAlaProPheSerCysCysValProAspProAlaGlnIleLysValIleAspThrGlnCys 200
Db      613 GGGGTCCCTTCTCTCTGCTGTGCGCAGATCTCTGCGCAAAAGTTGTGAACACACAGAGT 672
Qy      201 GlyTyrAspValArgIleGlnLeuLysSerIleTyrAspGluSerIlePheThrIleGly 220
Db      673 GGATATGATGTCAGAGTTCAGTGAAGAGCAGTGGATGATCATCTTCACGAAAGGC 732
Qy      221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
Db      733 TGCATCCAGGCGCTGCAAAAGCTGTGCTCCCGCAACATTTACATTTGTGGTGGCTCTTC 792
Qy      241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
Db      793 ATGGCATCTGCTGCTGTGAGATATTTGGCATCTTCTGCGCAAGACGCTGATCTCGAC 852
Qy      261 IleGluAlaValIleLysAlaGlyHisHisPhe 270
Db      853 ATCGAGCGAGTCAGAGCCGCCCATCATCTTC 882

RESULT 3
LOCUS   HSM801611      2556 bp      mRNA      linear      PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp564B1037 (from clone DKFZp564B1037);
complete cds.
ACCESSION AL136638
VERSION   AL136638.1  GI:12052801
KEYWORDS .
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2556)
AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE      Direct Submission
JOURNAL    Submitted (10-JUL-2002) GSF, Institut fuer Bioinformatik MIPS,
           Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany
           Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
           Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
           Heidelberg/Germany) within the cDNA sequencing consortium of the
           German Genome Project.
COMMENT   This clone (DKFZp564B1037) is available at the RZPD in Berlin.
           Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
           Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
           information about the clone and the sequencing project is available
           at http://mips.gsf.de/proj/cDNA/.

```

```

FEATURES
    source          location/Qualifiers
                   1..2556
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /chromosome="10"
                   /clone="DKFZp564B1037"
                   /tissue_type="brain"
                   /clone_lib="564 (synonym: hfdr2). Vector pAMP1; host
                   X1-2bui; sites NotI + SalI"
                   /dev_stage="fetal"
gene              1..2556
                   /gene="DKFZp564B1037"
                   /note="strong similarity to tetraspan NET-4"
CDS               79..891
                   /gene="DKFZp564B1037"
                   /codon_start=1
                   /product="hypothetical protein"
                   /protein_id="CAB66573.1"
                   /db_xref="GI:12052802"
                   /db_xref="SPTREMBL:O9H0U1"
                   /translation="MAYRYRSNAKVSQWKYLLEFSNIIEMILAGVFLGVGLMASEK
                   GVLSDTIKYTRRMGIDPVYLVLMGVGMFTLPAGCVGLARENICLINFECTIYLIE
                   FLELAVAVLAFLEQDWVRDREFFREFSNITSKSRDDIDLQNVNIDSLOKANQCCGAYPE
                   DMDLVNPFNCSGASYSREKGVPEFSCVDPAPQKVNVTGCGYDVRITQLSKWDSEIFT
                   KGCIOALESWLPNRIYIVAGVFALISLIQIFGLFARLISDIENAVKAGRHP"
polyA_site        2546
                   /gene="DKFZp564B1037"
BASE COUNT       500 a 729 c 683 g 644 t
ORIGIN
Alignment Scores:
Pred. No.:        2,45e-152      Length:      2556
Score:            1438.00        Matches:     268
Percent Similarity: 99.63%       Conservative: 1
Best Local Similarity: 99.26%    Mismatches:   1
Query Match:      99.24%         Indels:       0
DB:               Gaps:          0

US-09-905-674-2 (1-270) x HSM801611 (1-2556)
Qy      1 MethIstYrYrArGyYrSerAsnAlaLysValSerCysTrpTyrIstYrIleuLeuPhe 20
Db      79 ATGCACATATATATAGATACCTCTAACGCCAGGTGAGTGTGTGATACCAAGTACCTCTTTC 138
Qy      21 SerTyrAsnIleIlePheTrpLeuAlaGlyValPheLeuGlyValGlyLeuTrpAla 40
Db      139 AGCTACCAACATCATCTCTGTGTGCTGTGAGTGTCTCTCTGAGAGCGGGCTGTGGCCA 198
Qy      41 TrpSerGluLysGlyValIleuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
Db      199 TGGAGCGAAAAGGCTGTGCTGTCCGACCTCACCAAAATGACCCGGATGCTGGAATTCGAC 258
Qy      61 ProValValIleuValIleuMetValGlyValIleuPheThrLeuGlyPheAlaGlyCys 80
Db      259 CCTGTGTGTGCTGTGCTGTGATGTGGCGGTGTGTATGTTTACCCCTGGGGCTTCGCGGCTGC 318
Qy      81 ValGlyAlaLeuAurGluAsnIleCysLeuLeuAsnPhenPheCysGlyThrIleValLeu 100
Db      319 GTGGGGGCTCTCGGGAGAAATATGTGCTTCAACTTTTCTGTGGCACCATTCGTGCTC 378
Qy      101 IlePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArg 120
Db      379 ATCTTCTTCTCTGAGCTGTGCTGTGGCCGTGCTTCTGTTCCAGGACTGGGTGAGG 438
Qy      121 AspArgPheArgGluPhePheGluSerAsnIleCysSerTyrArgAspAspIleAspLeu 140
Db      439 GACCGGTTCCGGAGTTCTTCAGAGCAACATCAAGTCTTACCCGGAGCATATCGATCTG 498
Qy      141 GlnAsnLeuIleAspSerLeuGlnIleLysAlaAsnGlnCysGlyAlaIleTyrGlyProGlu 160
Db      499 CAAACCTCATCGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGGCCCTGAA 558
Qy      161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180

```

|||||
Db 559 GACTGGAGCCTCAACGCTCACTCAATGACAGCGGTGCCAGACAGCCGAGAGAGATGTC 618
|||
Qy 181 G1YVALPropheserCysValProAspProAlaGlnValValAsnThrGlnCys 200
|||
Db 619 GGGGTCCCTTCCTCCGCTGCGTGCAGATCCCTGCGCAAAAAGTTGTGAACACAGCTGT 678
|||
Qy 201 G1TYrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
|||
Db 679 GGATATGATGTACAGATTGCTGGAAGACAGTGGGATGATGCTTCATCTTCACGAAAGGC 738
|||
Qy 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPhe 240
|||
Db 739 TGCATCCAGCGCTGGAAGAGCTGCTCCGCGAACAATTACATGTGCGTGCCTCTTC 798
|||
Qy 241 ILeAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
|||
Db 799 ATGCCATCCGCTGTGTGCGATATTGGCATCTTCCTGCGCAAGACGCTGATCTCAGAC 858
|||
Qy 261 ILeGluAlaValLysAlaGlyHisHisPhe 270
|||
Db 859 ATCGAGGCACTGAAGCCGCGCTCCTCCTTC 888
|||
RESULT 4
BC025568 2428 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, similar to region containing hypothetical protein
DEFINITION MGCL1352; slingshot, clone MGC:36595 IMAGE:5322531, mRNA, complete
cds:
ACCESSION BC025568
VERSION BC025568.1 GI:19343883
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2428)
REFERENCE Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgcen@rl.nih.gov
Blakesley, N., Aylee, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stanlipop, S., Thomas, P.J., Touchman, J.W.,
Tsuang, C., Vogt, J.L., Walker, M.A., Weherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNU at: <http://image.llnl.gov>
Series: IRAN Plate: 56 Row: J Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction.
FEATURES
Source Location/Qualifiers
1..2428
/organism="Mus musculus"
/db_xref="taxon:10090"

/map="FVB/N"
/clone="MGC:36595 IMAGE:5322531"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_id="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
64..876
/codon_start=1
/product="Similar to region containing hypothetical
protein MGCL1352; slingshot"
/protein_id="AAH25568.1"
/db_xref="GI:19343884"
/translation="MHYRYSNAEVSQWYKLFSEYIVEMLAGVVELGGLAMSEK
GVSLDKTRRLHGIDPVYIVLWGVMMIFRAGCGVALREMICLKEFCAIVLIF
FLFLAVAVLFLQDWRDRFRFEFESNLSKSYRDDIDLNLDSLOKAOCCGATPE
DWDLNVFNCGSASYSERKGVPSCCVPAPAKVNVTCGDVRIQLSKMDFLET
KGCIOALEGWLPRNIYIVAGVFVIAISLQIFGIPLARTLISDEAVKAGHHF"
BASE COUNT 517 a 640 c 624 g 647 t
ORIGIN
Alignment Scores:
Pred. No.: 2,48e-150 Length: 2428
Score: 1420.00 Matches: 263
Percent Similarity: 98.52% Conservative: 3
Best Local Similarity: 97.41% Mismatches: 4
Query Match: 98.00% Indels: 0
DB: 10 Gaps: 0
US-09-905-674-2 (1-270) x BC025568 (1-2428)
Qy 1 MethISTYrYrArgTyrSerAsnAlaLysValSerCysTrpYrLysTyrLeuPhe 20
|||||
Db 64 ATGCACCTTTATGATCTCGACGCCGAGGTGACGCTGCTGATCAAGTACTCTCTTC 123
|||||
Qy 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40
|||||
Db 124 AGCTACATATTCGCTCTTTGGCTGGTGAAGTGTCTCTCTTGAGTGGCGGCTGTGGCA 183
|||||
Qy 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
|||||
Db 184 TGGAGCGAAAGGGGTGCTGCTCGACCTCACCAAGGTGACCCGCTTGCAATGTAC 243
|||||
Qy 61 ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80
|||||
Db 244 CCCGCGCGCTGCTGTGATGTGTGGCGCTGTGTATGTTACACACGCGGATTCGACGCTGT 303
|||||
Qy 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
|||||
Db 304 GTGGGGCGCTCCGAGAGAACATCGCTGCTCAAGTTTCTGTGGGCGCATGTGCTCTC 363
|||||
Qy 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAsnTrpValArg 120
|||||
Db 364 ATCTTCTTCCTGGAACGTGGCGCTGCTGTGGCTTTTATVTCGAAGACTGGGTAGA 423
|||||
Qy 121 AspArgPheArgGluPhePheGlnSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
|||||
Db 424 GACCGGTTCCGGGAATTCCTTGACAGCAACATCAAGTCTATGCGGATGATCATGACCTG 483
|||||
Qy 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
|||||
Db 484 CAGAACCTCATTTGACTCCCTTCAGAAAGCTATATGATGCTGGGGGCTTCACGGCCCTGAA 543
|||||
Qy 161 AspTrpAspLeuAsnValTyrPheAsnGlnSerLysValAserTyrSerArgGluLysCys 180
|||||
Db 544 GACTGGGACCTCAATGCTACTTCACTGCAAGTGTGCCAGCTACACCCAGGAAATGT 603
|||||
Qy 181 GlyValPropheserCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
|||||
Db 604 GGGGTACCTTCTCTGCTGTGTGCGATCTGCAAAAAGTCGTGAACACACAGCTGT 663
|||||
Qy 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
|||||

Db 664 GGCTATGATGTCGGATTCAGTGAAGACCAAGGGATGAGTTCATCTTACAAAAGGA 723

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
|||||
Db 724 TGCATCCAGGCTCTGGAAAGCGTGGCTGCCAGAACATCTACATGTGTGCTGTCTTC 783

QY 241 ILeAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
|||||
Db 784 ATGGCATCTCAGCTGTCAGATTTTGGCATCTCTCGAGAGACCCGTGATCTCAGAC 843

QY 261 ILeGlnAlaValIleAlaGlyHisHisPhe 270
|||||
Db 844 ATCGAGGACGATGAAGGACGACATCATCTTC 873

RESULT 5
BC024611

LOCUS BC024611 2498 bp mRNA linear ROD 07-AUG-2002

DEFINITION Mus musculus, similar to region containing hypothetical protein MGCI1352; slingshot, clone MGC:28503 IMAGE:4188261, mRNA, complete cds.

ACCESSION BC024611 GI:19354053

VERSION BC024611.1

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2498)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarune, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 37 Row: e Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.
Location/Qualifiers

FEATURES
source 1..2498
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:28503 IMAGE:4188261"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_lib="NCL CGAP_SG2"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 55..867
 /codon_start=1
 /product="Similar to region containing hypothetical protein MGCI1352; slingshot"
 /protein_id="AAH24611.1"
 /db_xref="GI:19354054"
 /translation="MHYIRISNAEYSCWKYKLFESYIVFWMAGVFLGGLMAWSK
GVLSDLRKIVRLHGIIDPVILWGVVMTLGFAGVGALRENIICLLKFCGAILLIF

CDS

FILELAVAVIAFILEQDWRDRFREFEFESNIKSYRDIIDLONLIDSLKANOCGAYGPE
DMDLNVFNCSSGASYSREKQYFSCCVDPAPKVVNCTCGVDVRLSKSMDERIFT
KGCIOLEGLPKNIRIYVAGVFALISLDIFGIFLARILISIDIEAVKAGHNE"

BASE COUNT 542 a 644 c 631 g 681 t

ORIGIN

Alignment Scores:
Pred. No.: 2,58e-150 Length: 2498
Score: 1420.00 Matches: 263
Percent Similarity: 98.52% Conservative: 3
Best Local Similarity: 97.41% Mismatches: 4
Query Match: 98.00% Indels: 0
Gaps: 0

US-09-905-674-2 (1-270) x BC024611 (1-2498)

QY 1 MethIstYrYrArGtYrSerAsnAlaLysValSerCysTrpYrIstYrLeuLeuPhe 20
|||||
Db 55 ATGCACATATATATGATCTCGACGCGAGGTGAGTGTCTTCTTGAGAGTGGGCTGGCA 114

QY 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValIstYrLeuTrpAla 40
|||||
Db 115 AGCTACATATCTCTTTTGGCTGGCTGGAGTGTCTTCTTGAGAGTGGGCTGGCA 174

QY 41 TrpSerGlyIstYrValLeuSerAspLeuThrIstYrValThrArgMetHisGlyIleAsp 60
|||||
Db 175 TGGAGCGAAAGGATGTGTCTCTCGACCTACCAAGGTGACCCGGTTGCATGGAATCGAC 234

QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
|||||
Db 235 CCCGTCGGCTGGCTGTGATGTGGCTGGCTGGCTGTGATGTGATGTTACACATCGGATTCGAGGCTGT 294

QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPheCysGlyThrIleValLeu 100
|||||
Db 295 GTCCGGGCGCTCCGAGGAACATCTGCTGCTCAAGTTTCTGTGGCCATTGTCTC 354

QY 101 IlePhePheLeuGlnLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120
|||||
Db 355 ATCTTCTTCCGGAACGCGCGTGGCTGTGGCTTTTATTTCAGAGCTGGTGAGA 414

QY 121 AspArgPheArgGlyPhePheGlnSerAsnIleLysSerTyrArgAspIleAspLeu 140
|||||
Db 415 GACCGGTCCGGAATCTTCGAGAGCAACATCAATCGGATGATCGACCTG 474

QY 141 GlnAsnLeuIleAspSerLeuGlnIstYrAlaAsnGlnCysGlyAlaTyrGlyProGln 160
|||||
Db 475 CAGAACCTCATTTGACTCTCTCAGAAAGCTAATCAGTGTGCGGGCTTAAGGCCCTGAA 534

QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyIstYrCys 180
|||||
Db 535 GACGCGGACCTCATGTCTTCTTCACTGACAGTGTGCCAGCATACAGCCGAGAAATGT 594

QY 181 GlyValProPheSerCysCysValProAspProAlaGlnIstYrValAlaAsnThrGlnCys 200
|||||
Db 595 GGGGTACCTCTCTCTCTCTGTCGACATCTGCAAAAATCTCGAACACACAGTGT 654

QY 201 GlyTyrAspValArgIleGlnIleuLysSerIstYrAspIstYrLeuIlePheTrpIstYrGly 220
|||||
Db 655 GGGTATGATGTCCGATTCAGCTGAAAGCAAGGGGATAGTTCACTTTACAAAAGGA 714

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
|||||
Db 715 TGCATCCAGGCTCTGGAAAGCGTGGCTGCCAGAACATCTACATGTGTGCTGTCTTC 774

QY 241 ILeAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
|||||
Db 775 ATTCGCATCTCAGCTGCGAGATTTTGGCATCTCTCGCGAGGACCCGTGATCTCAGAC 834

QY 261 ILeGlnAlaValIstYrAlaGlyHisHisPhe 270
|||||
Db 835 ATCGAGGACGATGAAGGACGACATCATCTTC 864

RESULT 6

BC026574 2500 bp mRNA Linear ROD 07-AUG-2002
 LOCUS BC026574 Mus musculus, clone MGC:36554 IMAGE:4954874, mRNA, complete cds.
 DEFINITION BC026574
 ACCESSION BC026574.1 GI:20072352
 VERSION BC026574.1
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2500)
 REFERENCE Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxll.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plates: 59 Row: a Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.
 FEATURES
 source Location/Qualifiers
 1. 2500
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:36554 IMAGE:4954874"
 /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 50. 862
 /codon_start=1
 /product="Unknown (protein for MGC:36554)"
 /protein_id="AAH26574.1"
 /db_xref="GI:20072353"
 /translation="MHYRYSNAEYSCYKYLFSYNIVFVLGAVFLGVLAMSEK GVLSDLTAVTRLHGIDPVVLVLMGVVMTLGFAGCVGLRENICLKFPGCAVLIF FLELAVAVLAFIFODMVRDRREFESNISKYRDIDQLNIDSLQKANOCCGAYGPE DMIDLNVYFNGSGASYSREKCGPSCCVDPKAVKVNOCGYDRIOLKSWDEFIIT KGCIOALSGMLPRNIYIVAGFVAFISLIQIGIFLARIILSIDIAVVAAGHHF"
 CDS
 545 a 643 c 631 g 681 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,58e-150 Length: 2500
 Score: 1420.00 Matches: 263
 Percent Similarity: 98.52% Conservative: 3
 Best Local Similarity: 97.41% Mismatches: 4
 Query Match: 98.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-905-674-2 (1-270) x BC026574 (1-2500)
 QY 1 MethisTyrTyrArgTyrSerAsnAlaLysValSerCysTrrPrYrIlystYrIleuLeuPhe 20
 DB 50 ATGCACATATTATAGATACGGAACGCCGAGTCAGCTGCTGTACAAAGTACCTGCTCTTC 109

QY 21 SerTyrAsnIleIlePheTriPLeuAlaGlyValValPheLeuGlyValGlyLeuTriPala 40
 DB 110 AGCTACATATATGCTCTTTTGCTGGCTGGAGTTCTCTTCTTGAGTCGGGGCTGGGCA 169
 QY 41 TrpSerGluYsglyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
 DB 170 TGGAGCGMAAAGGGTGTGTGTGCTCCGACCTCACCAAGGTGACCCGGTTGCATGTGAATCGAC 229
 QY 61 ProValValLeuValLeuMetValGlyValMetPheThrIleuGlyPheAlaGlyCys 80
 DB 230 CCCGTGCTGCTGCTGTGAATGCTGGGGCTGTGTATGTCACACTGGGATTCGCAAGGCTGT 289
 QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 DB 290 GTCGGGGCCCTCCGAGAGAACATCTGCTGCTCAAGTTTCTGTGGGCCATTGTGCTC 349
 QY 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTriPValArg 120
 DB 350 ATCTTCTTCTCGAAACTGGCCGCTGGCGCTGTGGCTTTTATTCMAAGACTGGTGAGA 409
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
 DB 410 GACCGGTTCGGGAATCTTCGAGAGCAACATCAAGTCATGCGGATGACATCGACCTG 469
 QY 141 GluAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160
 DB 470 CAGAACCTCATGTACTCCCTTCAGAAAGCATATCATAGTGTGCGGGCTTCAGGCCCTCGAA 529
 QY 161 AspTriPAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCys 180
 DB 530 GACTGGGACCTCATGCTACTTCACTGACAGTGTGCTCCAGCTACACCGAGAAATGT 589
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
 DB 590 GGGGTACCTTCCTCTGCTGTGTCAGATCTGCAAAAAGCTGGAACACACACTGT 649
 QY 201 GlyTyrAspValArgIleGlnLeuLysSerTyrTrpAspGluSerIlePheThrLysGly 220
 DB 650 GCGTATGATGTCGCCGATTCACGTAAGAGCAAGTGGATGATGATCATCTTTACAAAAGCA 709
 QY 221 CysIleGlnAlaLeuGluSerTriPLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
 DB 710 TGCAATCCAGCGCTCGAAGGCTGGCTGCCAGCAACATCATATTTGGCTGTGCTCTTC 769
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrIleLysSerAsp 260
 DB 770 ATTCGCATCTCAGCTGTCAGATTTTGGCATCTTCTGTGGAGAGCCGTGATCTCGAC 829
 QY 261 IleGluAlaLysAlaGlyHisHisPhe 270
 DB 830 ATCGAGCGACGTGAAGCGCAGGCATCACTTC 859
 RESULT 7
 LOCUS BC002920 2502 bp mRNA Linear PRI 12-JUL-2001
 DEFINITION BC002920 Homo sapiens, similar to transmembrane 4 superfamily member 9,
 clone MGC:11352 IMAGE:3954042, mRNA, complete cds.
 ACCESSION BC002920 GI:12804132
 VERSION BC002920
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 2502)
 REFERENCE Submitted (05-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK

COMMENT

Contact: MGC help desk
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC/DCRD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNT)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/nisc.mcq@nigrl.nih.gov>
Contact: Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Kalins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantiripop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNT at: <http://image.llnl.gov>
Series: IRAL Plate: 15 Row: e Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
source

1..2502

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:11352 IMAGE:3954042"
/tissue_type="skin, melanotic melanoma."
/clone_lib="NIR_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
63..824
/codon_start=1
/product="similar to transmembrane 4 superfamily member 9"
/protein_id="AAH02920.1"
/db_xref="GI:12804133."
/translation="MHYRYSNAKYSCWKYLLFSYNIIFWGVSLDLTKVTRMGIDP
VVLVLMGVVMTGLFAGCVGALRENICLNFECTIVLIFLELAVAVLAFPODMV
RDRREFESNSKSYRDIDLONLIDSLOKACGACGEMDINLVNFCGASYSR
KCGVPEFCVCPDPAOKVYVNOGQYDVRLOKSKDESFPTGCIQALDSMLPRNIYI
VAGVFIAISLIDIFGLFARTISDIEAVKAGHNP"

CDS

BASE COUNT 505 a 711 c 661 g 625 t
ORIGIN

Alignment Scores:

Pred. No.: 4,18e-141 Length: 2502
Score: 1338.50 Matches: 253
Percent Similarity: 93.70% Conservative: 0
Best Local Similarity: 93.70% Mismatches: 0
Query Match: 92.37% Indels: 17
DB: 9 Gaps: 1

US-09-905-674-2 (1-270) x BC002920 (1-2502)

QY 1 Methistyrtyrargtyrserasnalaalysvalsercystirpyrtyrtyrleuaphe 20
|||||
Db 63 ATGCACATATATAGATCTCTAACGCCAGGTGAGTGGTGAACAAGTACCTCTTTC 122
QY 21 SerTyrAsnIleIlePheTTPLeuAlaGlyValValPheLeuGlyValGlyLeuTPaLa 40
|||||
Db 123 AGCTACAACATCACTCTCTGG----- 143
QY 41 TrpSerGluysGlyValLeuSerAspLeuThrIlyValThrArgMetHisGlyIleasp 60
|||||
Db 144 -----GGTGGCTGCTCCGACCTCAACCAAGTGACCCGGATGCATGCAATGCAC 191
QY 61 ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80
|||||
Db 192 CCTGTGGTGGCTGGATGATGGTGGGTGATGTTCACCGTGGGCTTCGCGGCTGC 251
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100

Db 252 GTGGGGGCTCGCGGAGAAATATCTGCTGTCACTTTTGTGTGGCACATCGGCTC 311
QY 101 IllePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120
Db 312 ATCTTCTCTCGTAGAGCTGGCTGGCGGCTGCTGCTCTCTGTTCCAGGACTGGGTGAGG 371
QY 121 AspArgPheArgGluPhePheGlnSerAsnIleLysSerTyrTrpAspPheIleAspLeu 140
Db 372 GACCGGTTCCGGGAGTTCTTTCAGAGCAACATCAAGTCTACCGGAGCATATGATCTG 431
QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
Db 432 CAAACCTCATCGACTCCCTTCACAAAGACATACAGAGCTGTGGCGCATATGGCCCTGAA 491
QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluCys 180
Db 492 GACTGGGACCTCAACGCTACTTCAATTGCAGCGCTCCAGCTACAGCCAGAGAAAGTGC 551
QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200
Db 552 GGGGTCCCTTCTCTCTGCTGGTCCAGATCTCGCAAAAAGTTGTAACACACAGCTGT 611
QY 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220
Db 612 GCATATGATGTACAGATTACACTGACAGCAAGTGGATGATGTCATCTTCACGAAAGGC 671
QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
Db 672 TGCAATCCAGGGCGCTGGAAGCTGGCTCCCGCGGAACATTTCATTTGGTGGCTGCTTC 731
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
Db 732 ATCCGCAATCTCGCTGTTCAGATATTTGGCATCTTCGCGAAGAGCTGATCTCAGAC 791
QY 261 IleGluAlaValLysAlaGlyHisHisPhe 270
Db 792 ATCGAGGCAGTGAAGCGCGGCCATCACTTC 821
RESULT 8
AF065389 1405 bp mRNA linear PRI 28-Apr-2000
LOCUS
DEFINITION Homo sapiens tetraspan NTR-4 mRNA, complete cds.
ACCESSION AF065389
VERSION AF065389.1 GI:3152702
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1405)
Serru, V., Dessen, P., Bouchaix, C. and Rubinstein, E.
Sequence and expression of seven new tetraspans
JOURNAL Biochim. Biophys. Acta 1478 (1), 159-163 (2000)
MEDLINE 20185353
PUBMED 10719184
2 (bases 1 to 1405)
Rubinstein, E., Serru, V. and Bouchaix, C.
Direct Submission
Submitted (14-MAY-1998) INSDERM 0268, 14 av Paul Vallant Couturier,
Villejuif 94807, France
FEATURES
source
1..1405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE ID 219547"
348..1154
/note="similar to Tspan-5; TM5F"
/codon_start=1
/product="tetraspan NTR-4"
/protein_id="AAC17120.1"
/db_xref="GI:3152703"
/translation="MSGKHKGPVSCIKYIFGFNVIFWFLGITFLGIGLMMANKE

Accession	Gene	Species	Length (bp)	Insertion (bp)	Source
U00001	phr	<i>Phlebotomus perniciosus</i>	45	0	GenBank
U00002	phr	<i>Phlebotomus perniciosus</i>	468	0	GenBank
U00003	phr	<i>Phlebotomus perniciosus</i>	65	0	GenBank
U00004	phr	<i>Phlebotomus perniciosus</i>	528	0	GenBank
U00005	phr	<i>Phlebotomus perniciosus</i>	85	0	GenBank
U00006	phr	<i>Phlebotomus perniciosus</i>	588	0	GenBank
U00007	phr	<i>Phlebotomus perniciosus</i>	105	0	GenBank
U00008	phr	<i>Phlebotomus perniciosus</i>	648	0	GenBank
U00009	phr	<i>Phlebotomus perniciosus</i>	125	0	GenBank
U00010	phr	<i>Phlebotomus perniciosus</i>	708	0	GenBank
U00011	phr	<i>Phlebotomus perniciosus</i>	145	0	GenBank
U00012	phr	<i>Phlebotomus perniciosus</i>	768	0	GenBank
U00013	phr	<i>Phlebotomus perniciosus</i>	165	0	GenBank
U00014	phr	<i>Phlebotomus perniciosus</i>	828	0	GenBank
U00015	phr	<i>Phlebotomus perniciosus</i>	185	0	GenBank
U00016	phr	<i>Phlebotomus perniciosus</i>	888	0	GenBank
U00017	phr	<i>Phlebotomus perniciosus</i>	205	0	GenBank
U00018	phr	<i>Phlebotomus perniciosus</i>	948	0	GenBank
U00019	phr	<i>Phlebotomus perniciosus</i>	225	0	GenBank
U00020	phr	<i>Phlebotomus perniciosus</i>	1008	0	GenBank
U00021	phr	<i>Phlebotomus perniciosus</i>	245	0	GenBank
U00022	phr	<i>Phlebotomus perniciosus</i>	1068	0	GenBank
U00023	phr	<i>Phlebotomus perniciosus</i>	265	0	GenBank
U00024	phr	<i>Phlebotomus perniciosus</i>	1128	0	GenBank
U00025	phr	<i>Phlebotomus perniciosus</i>	266	0	GenBank
U00026	phr	<i>Phlebotomus perniciosus</i>	1131	0	GenBank
U00027	phr	<i>Phlebotomus perniciosus</i>	10	0	GenBank
U00028	phr	<i>Phlebotomus perniciosus</i>	344	0	GenBank
U00029	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00030	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00031	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00032	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00033	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00034	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00035	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00036	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00037	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00038	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00039	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00040	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00041	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00042	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00043	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00044	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00045	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00046	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00047	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00048	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00049	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00050	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00051	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00052	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00053	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00054	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00055	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00056	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00057	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank
U00058	phr	<i>Phlebotomus perniciosus</i>	3175	0	GenBank</

FEATURES	source	Location/Qualifiers
gene	CDS	1..3175 /organism="Mus musculus" /strain="Swiss Webster/NIH" /db_xref="taxon:10090" 1..3175 /gene="Tspan5" 388..1194
BASE COUNT	667 a 811 c 840 g 857 t	
ORIGIN	1..08e-87 Score: 867.00 Percent Similarity: 76.25% Best Local Similarity: 59.00% Query Match: 59.83% DB: 10 Gaps: 0	
US-09-905-674-2 (1-270) x AF121344 (1-3175)		
QY	6	TyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuLeuPheSerTyrAsnIleLe 25
Db	403	TCAAGAGGCTCGAAGTACAGTGTGGATGCATCAAAATCACTCATTTTGGCTTCATGTCAFA 462
QY	26	PheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAlaTrpSerGly 45
Db	463	TTTTGGTTTTGGGATAACGTTTCTGGAAACGGACACTGGCGCGGATGCAAAAAGT 522
QY	46	ValLeuSerAspLeuThrLysValAlaPheGluHisGlyIleAspProValValLeuVal 65
Db	523	GTCTCTCCACATCTCGTCATCCACCGACCTGGGCTTTGACCCAGTGGCTTTTC 582
QY	66	LeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysValAlaLeuArg 85
Db	583	CTTGTCGTGGAGAGAGTATGTTCACTCTCGGGGTTTCCAGGGTGCATCGAGACATTCGG 642
QY	86	GluAsnIleCysLeuLeuAsnDhePheCysGlyThrIleValLeuIlePhePheLeuGlu 105
Db	643	GAAACACCTTCTCTCTCAAGTTTCTGTGTTCTGCTCCGCGGATATTCTTCTCGGA 702
QY	106	LeuAlaValAlaValLeuAlaPheLeuPheGluAsnTrpValAlaArgAspArgPheArgGlu 125
Db	703	CTCACTGCTGGGGTGTGGCATTTGTTTTCACAAAGCTGGATCAAGACCGTGAATTC 762
QY	126	PhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGluAsnLeuIleAsp 145
Db	763	TTTATTAAACAACAATCGACCTCAGACATGATGATCTACATCTACAGACCTCAATAGAC 822
QY	146	SerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsn 165
Db	823	TTACCCACAGATATTGCGAGTGGTGGGGCTTTTGGAGCTGATGATGGAACCTAAT 882
QY	166	ValTyrPheAsnCysSerCylAlaSerTyrSerArgGlnLysCysGlyValProPheSer 185
Db	883	ATTTACTTCAATTGACACAATTCACATGCAACCCAGACCGATGGGGTGCATATTC 942
QY	186	CysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArg 205
Db	943	TGCTGCACCTAAAGACCCCGGAGAAATGTCATCAACACTCACTAGTGGCTATGATCCAG 1002
QY	206	IleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGlyCysIleGlnAlaLeu 225

```
Db 1003 CAGAAACGAAAGTGCACAGATTGTAATCTACACAAAGGCTGTGCCCCAGTTT 1062
      ::::: ||:::|||||||:::
Qy 226 GUSerTPrleuprArgrsAnleTyrlleValAlaGlyValPheleAlaIleSerleu 245
      ||| ||||| ||::: |||||:::||||| ||::: |||||
Db 1063 GAGAAAGTGTGCTACAGACATTTAAACCATTCGTGGCTGATTTTATAGCATTCGATTG 1122
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 246 LeuGAllePheGlylePheleuAlaArgThrleuIleSerAspIleGluAlaValys 265
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1123 CTACAGATTTTGGATATGCGCTGACACAGAAATTGGTGTAGTACATTGAACCTGTCAAG 1182
      |||
Qy 266 Ala 266
      |||
Db 1183 GCT 1185
      |||

RESULT 11
LOCUS BC010346 1516 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, RIKEN cDNA 2210021G21 gene, clone MGC:6941
IMAGE:2811935, mRNA, complete cds.
ACCESSION BC010346
VERSION BC010346.1 GI:16307592
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1516)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgabs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarane, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 5 Row: f Column: 18.

FEATURES
    source
        1..1516
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /map="C57BL/6J"
            /clone="MGC:6941 IMAGE:2811935"
            /tissue_type="Mammary tumor. WAP-IGF alpha model. 7 months
            old, gross tissue."
            /clone_11b="NCI_CGAP_Mam5"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"
            136..948
            /codon_start=1
            /product="RIKEN cDNA 2210021G21 gene"
            /protein_id="AAH10346.1"
            /db_xref="GI:16307593"
            /db_xref="LocusID:74257"
            /translation="MPGKHQOFDPEVCGCGKFLFGFNIVFWYIALGIMAWG
            EKVLSNLSALTDGGLDPVWLFVYGVGMYSLVFGAGTGLARENTFLKRFYSFLGL
            IFFELAGITLAFVKDILRDQLNLFINNNAKAYRDDIDLNLIDFQEWSCCGGAS
            PDMWNLNTYFNCTDLNPISRCRGVGFSCVDPADVEDVLTNOCGYDIRLKLLEQGS1
```

```
YTKCGVGFPEKMLQDNLIVAGVLGIALQLIFGLCLLAONLVSIDIKAVRANW"
BASE COUNT 277 a 417 c 461 g 361 t
ORIGIN
Alignment Scores:
Pred. No.: 1,98e-87 Length: 1516
Score: 861.00 Matches: 151
Percent Similarity: 76.89% Conservative: 52
Best Local Similarity: 57.20% Mismatches: 61
Query Match: 59.42% Indels: 0
Gaps: 0
Db: 10

US-09-905-674-2 (1-270) x BC010346 (1-1516)
Qy 3 TyrTyraGtyrSerAsnAlaValSerCysTrpTyrlsTyrlleuPheSerTy 22
      ::::::::::::::: ||| ||| ||||| |||||
Db 148 CACGACATTCACAGACCTGAGGTGGCTGGCGGAAATCTCTGTTGGCTTC 207
      ::::::::::::::: ||| ||| ||||| |||||
Qy 23 AsnIleIlePheTrPleuAlaGlyValAlaPheleuGlyValGlyleuTrPaLatrpSer 42
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 208 AACATGCTCTCTGGTGTGAGAGCCCTGCTGCGCATTTGGCCTCTGGCGGT 267
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 43 GluIysGlyValLeuSerAspLeuThrIysValThrArgMetHisGlyIleAspProval 62
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 268 GAGAAAGGCGTCTCTCCACATCTCGCGCGCTGAGATCGGCGGCTTGAACCCGTG 327
      ||| ::::::::::||| ||||| ||||| ||||| ||||| |||||
Qy 63 ValIeuValIleMetValGlyValAlaMetPheThrleuGlyPheAlaGlyValGly 82
      ||| ::::::::::||| ||||| ||||| ||||| ||||| |||||
Db 328 TGGCTGTTGTGTGGTGGGAGTGCATGTCAGTGTGGCTTGGCGGCTGCATTTGGG 387
      ||| ::::::::::||| ||||| ||||| ||||| ||||| |||||
Qy 83 AlaIeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIlePhe 102
      ||| ||||| ||||| ||||| |||||
Db 388 GCCCTCCGGGAAACACCTCTCTCTCAAGTCTTCTCTGTCTTCCGCGGCTCATCTTC 447
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 103 PheIeuGluIleuAlaValAlaValIleuAlaPheIeuPheGluAsnTrpValArgAspArg 122
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 448 TTCTGTGAGCTGGCGGGGGGATCTGCTGCTGCTCAAGATTTGATCCGAGACCA 507
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 123 PheArgGluPhePheGluSerAsnIleIysSerTyraArgAspIleAspIleuGluAsn 142
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 508 CTTAATCTCTTCATCAACACAAATGTCAAAGCTCAACGGGATGATCTCGACTTCGAAC 567
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 143 LeuIleAspSerLeuGlnIysAlaAsnGlnCysGlyAlaTyrglyProGluAspTrp 162
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 568 CTTATTCAGCTTGTCTCAGAGATACTGTGCTGCTGTGTGCTCCGAGGCCCATGACTCG 627
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 163 AspIleuAsnValTyraPheAsnCysSerGlyAlaSerTyraArgGluIysCysGlyVal 182
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 628 AACCTCAATATCTACTCAACTGCATGCACTGTGAACCAAGCCGTGAGCCGTGTGGGGTG 687
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 183 ProPheSerCysValProAspProAlaGlnIysValAlaAsnThrGlnCysGlyTy 202
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 688 CCTTTTCTCTCTGTGTAGGAGCCAGCGGAAATGCTCTCAATATCCAGTGTGCTAT 747
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 203 AspValArgIleGlnIleuIysSerIysTrpAspGluSerIlePheThrIysGlyIle 222
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 748 GACATTCGACATCAACATGAGACTCGACAGAGGCTCCATCTACACCAAGGCTGTGTG 807
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 223 GlnAlaIeuGluSerTrPleuProArgAnleTyrlleValAlaGlyValPheIleAla 242
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 808 GGCCAGTTTGAAGATGCTACAGACAACTGATCTGTGCGCGGCTTGTGGGG 867
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 243 IleSerIleuGlnIlePheGlyIlePheleuAlaArgThrleuIleSerAspIleGlu 262
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 868 ATGCGCTCTCTCGATATTTGGCTCTGCTGTGCGCCAGAACTCGTAGTGCATCAAG 927
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Qy 263 AlaValIysAla 266
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||
Db 928 GCGGTGAAGGCC 939
      |||||:::|||||::: || ::||| ||| ::||| ||||| |||||

RESULT 12
AX420468 813 bp DNA linear PAT 18-JUN-2002
LOCUS
```



```

DEFINITION      Sequence 3 from Patent WO0216603.
ACCESSION       AX420468
VERSION         AX420468.1  GI:21524616
KEYWORDS
SOURCE         human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS        Leiby, K.R.
TITLE          23228, a human tetraspanin family member and uses thereof
JOURNAL        Patent: WO 0216603-A 3 28-FEB-2002;
                Millennium Pharmaceuticals, Inc. (US)
FEATURES
source         1..813
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT     151 a 233 c 235 g 194 t
ORIGIN
Alignment Scores:
Pred. No.:      1 94e-87      length:      813
Score:          858.00      Matches:      154
Percent Similarity: 75.76%      Conservative: 46
Best Local Similarity: 58.33%      Mismatches: 64
Query Match:    59.21%      Indels:      0
DB:             Gaps:      0

US-09-905-674-2 (1-270) x AX420468 (1-813)
QY      3 TyrTyrAgtYrSerAsnAlaLysValSerCysTrrPtyrLysTyrLeuLeuPheSerTyr 22
       ::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      13 CACCACACATTTCCAGAACCTGAGTGGCTGCTGGCGGGAATACCTCTGTTGGCTTC 72

QY      23 AsnIleIlePheThrPheuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
       |||||:::|||||::: ||:::|||||::: ||:::|||||::: ||:::|||||::: ||:::
Db      73 AACATTGCTTCTGGTGGTGGAGCCCTGTTCCTGGCTATCGGCTCTGGCGCTGGGGCT 132

QY      43 GluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal 62
       |||||:::|||||::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db      133 GAAAGGGCTTCTCTCGAACATCTCAGCGCTGACAGATCTGGAGGCGCTTGACCCCGTG 192

QY      63 ValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysValGly 82
       ||| :::::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      193 TGGCTGTTGGTAGTAGTGAGGCGCTCATGTGCGTCTGCGCTTGGCTGCGCATGGG 252

QY      83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe 102
       |||||:::|||||::: |||||::: |||||::: ::::: |||||::: |||||:::
Db      253 GCCCTCCGGAGAACACCTTCCTCGCTCAAGTTTCTCCGTCTCCTCGCTCATCTTC 312

QY      103 PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArgAspArg 122
       |||||:::|||||::: :::::|||||::: :::::|||||::: :::::|||||::: :::::
Db      313 TTCTCTGAGCTGGCAACAGGATCCTGGCTTGTCTTCAAGAGACTGGATTCGAGACCG 372

QY      123 PheArgGluPhePheGluSerAsnIleCysSerTyrArgAspAspIleAspLeuGluAsn 142
       ||| :::::|||||::: :::::|||||::: :::::|||||::: :::::|||||::: :::::
Db      373 CTCAACCTCTTCATCAACAACAACGTCAGGCGCTACCGGAGACGACATTGACCTCCAGAAC 432

QY      143 IleuIleAspSerLeuGlnIysAlaAsnGlnCysGlyAlaIleArgIleProGluAspTrp 162
       |||||:::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      433 CTGATTGACTTGGTCAGGAATACCTGCTCTGTGGAGCCCGGAGCCCAATATACCTGG 492

QY      163 AspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyVal 182
       :::::|||||::: :::::|||||::: :::::|||||::: :::::|||||::: :::::
Db      493 AACCTCAATATCTACTTCAACTGCACTGATGTAACCCGAGCGGAGGCGCTGCGGGGTG 552

QY      183 ProPheSerCysCysValProAspProAlaGlnIysValAlaAsnThrGlnCysGlyTyr 202
       |||||:::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      553 CCTTCTCTGCTGCTCAGGAGCCCTGCGGAGGATGCTCTCAACACCAAGTGGGCTAC 612

QY      203 AspValArgIleGluLeuLysSerLysTrpAspGluSerIlePheThrIleLysGlySile 222
       |||||:::|||||::: ::::: |||||::: |||||::: |||||::: |||||::: |||||
Db      613 GAGCTCGGCTCAAACTGGAGCTGGAGCAGCAGGCGCTTATCCACCAACCAAGGCTGCGTG 672

```

```

QY      223 GlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValAlaPheIleAla 242
       ||| ||||| ||||| ::::: |||||::: |||||::: |||||::: |||||:::
Db      673 GGCACGTTTGAAAGCTGTCGAGCAACCTGATTTGTGGCGGAGCTTCATGGCC 732

QY      243 IleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgTrpLeuIleSerAspIleGlu 262
       |||||:::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      733 ATCGCCCTTCCTCAATCTTTGGCATCTGGCTGGCCGAGAACCTGTGATGACATCAAG 792

QY      263 AlaValLysAla 266
       |||||:::|||||:::
Db      793 GCAGTGAAGGCC 804

RESULT 13
AX420466      3184 bp      DNA      linear      PAT 18-JUN-2002
LOCUS
DEFINITION   Sequence 1 from Patent WO0216603.
ACCESSION   AX420466
VERSION     AX420466.1  GI:21524614
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS      Leiby, K.R.
TITLE        23228, a human tetraspanin family member and uses thereof
JOURNAL      Patent: WO 0216603-A 1 28-FEB-2002;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES
source       1..3184
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD35489.1"
            /db_xref="GI:21524615"
            /translation="MGKKHGHFDEPEVCGCGKFLFEGNIVFWVGLALPLAIGLWANG
            EKGVLNLSIALTDIGGLDPWLFVVGVSIVGFGCGIALNENFLKTFESVFLGL
            IFLELATGTGILAEVFKDMIRDQNLNLFNNNVKVARDDIDQNLIDYRAQEWKSCGARG
            PNDNLNIVFNCTDLNPSRERCGVFSCCVDRPAEDVILNOCGYDRALKLELPOQGF
            HTKCCVQGFKKMLQDNLIIVAGVPMGIALQIHGICLAQNLVSIDIAVAVANM"

BASE COUNT   675 a 883 c 945 g 681 t
ORIGIN
Alignment Scores:
Pred. No.:      1.13e-86      length:      3184
Score:          858.00      Matches:      154
Percent Similarity: 75.76%      Conservative: 46
Best Local Similarity: 58.33%      Mismatches: 64
Query Match:    59.21%      Indels:      0
DB:             Gaps:      0

US-09-905-674-2 (1-270) x AX420466 (1-3184)
QY      3 TyrTyrAgtYrSerAsnAlaLysValSerCysTrrPtyrLysTyrLeuLeuPheSerTyr 22
       ::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      180 CACCACATTTCCAGAACCTGAGTGGCTGCGCGGGAATACCTCTGTTGGCTTC 239

QY      23 AsnIleIlePheThrPheuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
       |||||:::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      240 AACATTGCTTCTGGTGGTGGAGCCCTGCTCTGCTATCGGCTCTGGGCTGGGCT 299

QY      43 GluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal 62
       |||||:::|||||::: :::::|||||::: :::::|||||::: :::::|||||::: :::::
Db      300 GAGAAAGGCGTTCTCTCGAACATCTCAGCGCTGACATCTGGAGAGCGCTTGACCCCGTG 359

QY      63 ValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysValGly 82
       ||| :::::|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db      360 TGGCTGTTGGTAGTAGTGAGGCGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 419

QY      83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe 102

```

```

Db 420 GGCCTCGGGAGAACACCTTCTGCTCAAGTTTTCGCGTGTCTGCTGCTCATCTTC 479
      |||||||
Qy 103 PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAspArg 122
      |||||||
Db 480 TTCTCTGGAGCTGGCAACAGGATCTCGGCTTGTCTTCAAGAGCTGATGAGACACAG 539
      |||||||
Qy 123 PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn 142
      |||
Db 540 CTCACCTCTTCTATCAACACACACACGATGAGCTTACCGGAGACATTTGACCTCCAGAAC 599
      |||
Qy 143 LeuIleAspSerLeuGlnAlaValAlaValLeuAlaValLeuAlaValLeuAlaValLeuAlaVal 162
      |||||||
Db 600 CTCATTGACTTTGGTGCAGAAATACAGTCTGTCTGTGGAGCCCGAGGCCCAATGACGTGG 659
      |||||||
Qy 163 AspLeuAsnValTyrPheAsnLysSerGlyAlaSerTyrSerArgGluLysCysGlyVal 182
      |||||||
Db 660 AACCTCAATATCTACTTCACTCACTGACCTGAACCCCGAGCCGAGCGGTGGGTG 719
      |||
Qy 183 ProPheSerCysCysValAlaProAspProAlaGlnLysValAlaValLeuAlaValLeuAlaVal 202
      |||||||
Db 720 CCCTTCTCTGCTGCTGCTGAGGACCCCTGGGAGGATGTCTTCAACACCAAGTGTGGTAC 779
      |||||||
Qy 203 AspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIle 222
      |||||||
Db 780 GAGCTCGGCTCAAACTGAGCTGAGGACGAGCAGCAGGCGTTCATCACACCAAGCTGTGGTG 839
      |||
Qy 223 GlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPheIleAla 242
      |||
Db 840 GGCCACTTTGAGCAAGTGGCTGAGGACGACACGATTTGTGTGGCGGACTTCTTCAATGGGC 899
      |||
Qy 243 IleSerLeuGlnIlePheGlyIlePheLeuAlaArgThrIleIleSerAspIleGlu 262
      |||||||
Db 900 ATGCGCCCTCTCAGATCTTTGGCATCTGCTGCGCCACAACCTGCTGATGATCAACAG 959
      |||
Qy 263 AlaValLysAla 266
      |||||||
Db 960 GCAGTGAAGCC 971
      |||
RESULT 14
AF053455 1408 bp DNA linear PRI 03-NOV-1998
LOCUS Homo sapiens tetraspan TM4SF (TSPAN-5) gene, complete cds.
DEFINITION AF053455
ACCESSION AF053455
VERSION AF053455.1 GI:2995864
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1408)
TITLE Sequences and expression of six new members of the
JOURNAL tetraspanin/TM4SF family
MEDLINE Blochim. Biophys. Acta 1399 (1), 101-104 (1998)
PUBMED 9714763
AUTHORS Todd,S.C., Doctor,V.S. and Levy,S.
TITLE Sequences and expression of six new members of the
JOURNAL tetraspanin/TM4SF family
AUTHORS Todd,S.C., Doctor,V.S. and Levy,S.
SUBMITTER Submitted (12-MAR-1998) Medicine, Stanford, 300 Pasteur Dr,
Stanford, CA 94305, USA
FEATURES
location/Qualifiers
1..1408
/organism="Homo sapiens"
/db_xref="taxon:9606"
352..1146
/gene="TSPAN-5"
352..1146
/gene="TSPAN-5"
/codon_start=1
/product="tetraspan TM4SF"

```

```

/protein_id="AAC69712.1"
/db_xref="GI:2995865"
/translation="MSGKHKKGEVSCCIKYPILFPGNVTFMFGLITPLGIGLWANNK
GVLSNISITTDLGGEPVWHLFLVYGVSMVILIFACIGALRENTLSVIGITFLEL
TAVLAFVFRDWIKDLYFFINNNIRATKDDIDLOLIDFTPEYMOCCARFADNML
NIYFNCTSDNASRERGVPSCTCKPADVDINTQGDARQKPEVDQIVITYGCV
PQEFKMIQNLNLTVAIGIFGIALQIFGICLQNLNLYSDIEAVASW"
BASE COUNT 309 a 354 c 384 g 345 t 16 others
ORIGIN
Alignment Scores:
Pred. No.: 4.09e-86 Length: 1408
Score: 849.00 Matches: 153
Percent Similarity: 75.86% Conservative: 45
Best Local Similarity: 58.62% Mismatches: 59
Query Match: 58.59% Indels: 4
DB: Gaps: 2
US-09-905-674-2 (1-270) x AF053455 (1-1408)
Qy 6 TyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuPheSerTyrAsnIleIle 25
      |||
Db 367 TACAAGGCTCTGAAGTCAGTGTGTCATCAATACTTCAATTTGGCTTCAATGTGCAT 426
      |||
Qy 26 PheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSerGluLysGly 45
      |||||||
Db 427 TTTTGGTTTGGGAAATACATTTCTTGGAATTGAGCTGTGGCATGAAATGAAGAA 486
      |||
Qy 46 ValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProValIleLeuVal 65
      |||||||
Db 487 GTTCTGTCCAACTCTTCATCCATCCAGATCTCGCGGCTTGTGACCAAGTTTGCTCTTC 546
      |||
Qy 66 LeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArg 85
      |||
Db 547 CTTGTGGTGGAGAGAGATGATTTATTTGGGATTTTCAGGATTCATGAGCGCTACGG 606
      |||
Qy 86 GluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlu 105
      |||
Db 607 GAAAC--ACTTCTCTTCTGTGCTCTGGA-----ATTATTTCTTCTCTGAG 654
      |||
Qy 106 LeuAlaValAlaValIleAlaPheLeuPheGlnAspTrpValAlaAspArgPheArgGlu 125
      |||
Db 655 CTCACCTGCGGAGTTCAGCATTTGTTTCAAAACGTGATCAAAAGCAGCTGTATTTC 714
      |||
Qy 126 PhePheGluSerAsnIleLysSerTyrArgAspIleAspLeuGlnAsnLeuIleAsp 145
      |||
Db 715 TTTATTAACACACACATCAGCATATCGGATGACATTTGATTTGCAAAACCTCATAGAC 774
      |||
Qy 146 SerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsn 165
      |||
Db 775 TTACCCAGCAAAATATTGGCAGTGTGGGCTTTTGGAGCTGATGTGGAACCTTAAT 834
      |||
Qy 166 ValTyrPheAsnLysSerGlyAlaSerTyrSerArgGluLysCysGlyValAlaProPheSer 185
      |||
Db 835 ATTACTTCATATTCACACGATTTCCAAATGCAAGTGCAGAGCGATGTGCTTCCATTTC 894
      |||
Qy 186 CysCysValAlaProAspProAlaGlnLysValAlaValAsnThrGlnCysGlyTyrAspValArg 205
      |||
Db 895 TGCTGCACATAAGATCCCGCAGAGATGTCTATCAACACTGATGTGCTATGATGCCAGG 954
      |||
Qy 206 IleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeu 225
      |||
Db 955 CAAAACCAAGAAATTGACACACAGATTTGATATTCACAGAAAGCGTGTGCCCACTTT 1014
      |||
Qy 226 GluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeu 245
      |||
Db 1015 GAGAGTGTGTGACAGCAAAATTTAAACCATCTGCTGTATTTTCATGAGCATTTGCA 1074
      |||
Qy 246 LeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGluAlaValLys 265
      |||
Db 1075 CTGCAGATATTGGGATATGCTGTGGCCAGAAATTTGGTTAGCGATATCGAAGCTGTACG 1134
      |||
Qy 266 Ala 266

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 19, 2002, 06:14:57 : Search time 2003 Seconds

(without alignments)
2183.116 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYSNAKVCWKYLLF.....IFLARTLISDEIAVKAGHHF 270

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh

-MODE=frame+pn.model
-Q=/cgn2_1/USPTO/spool/US0905674/runat_16122002.111846_6769/app-query.fasta.1.455
-DB=EST-QFMT=fastlap-SUPEX-first-MINMATCH=0.1-LOOPEXT=0
-UNITs-bits-STAR=1-END=1-MATRIX=blsum62-TRANS=human40.cdi-LIST=45
-DOCALLIGN=200-THR SCORE=Pct-THR MAX=100-THR MIN=0-ALIGN=15-MODE=LOCAL
-OUTFMT=ptio-NORM-ext-HEADSIZE=500-MINLEN=0-MAXLEN=2000000000
-USER=US0905674.ecgn.1.1.899-@runat_16122002.111846_6769-NCPU=6-ICPU=3
-NO_XLPHY-NO_MAP-LARGEQUERY-NEG_SCORES=0-WAIT-LONGLOG-DEV-TIMEOUT=120
-WARR-TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5-FGAPOP=6-FGAPEXT=7
-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1436	99.1	899	14	BQ212057
2	1430	98.7	1104	14	BM923308
3	1383	95.4	866	12	BG674208
4	1324	91.4	1125	14	BQ067138
5	1304	90.0	883	12	BG753541
6	1287.5	88.9	857	13	B1835459
7	1286	88.8	836	13	B1905852
8	1276	88.1	831	13	B1663945
9	1275	88.0	742	14	BQ572112
10	1271	87.7	951	9	AL519790
11	1246.5	86.0	797	13	BG913227
12	1246.5	86.0	870	14	BQ948800
13	1244	85.9	755	12	BG765719
14	1228	84.7	810	12	BG760633
15	1223	84.4	770	13	BG974843
16	1218	84.1	940	14	BQ943303
17	1197.5	82.6	845	12	BG862723
18	1197.5	82.6	764	13	B1853594
19	1181.5	81.5	1013	13	BG915134
20	1173	81.0	724	12	BG770453
21	1171	80.8	730	13	B151673
22	1165	80.4	728	13	B1556034
23	1156	79.8	824	13	B1912573
24	1144	79.0	925	9	AL539537
25	1133	78.2	970	12	BG760833
26	1128	77.8	694	13	BM051048
27	1122	77.4	751	12	BG246551
28	1120	77.3	841	12	BF026182
29	1118.5	77.2	924	12	BE895131
30	1090	75.2	736	13	B1251280
31	1088	75.1	1001	12	BF027356
32	1083	74.7	975	12	BF785494
33	1081	74.6	973	12	BG479961
34	1070	73.8	794	12	BF537170
35	1055.5	72.8	981	12	BG176342
36	1050.5	72.5	855	12	BF099030
37	1047	72.3	744	10	BE276487
38	1045	72.1	710	10	BE276329
39	1042	71.9	743	10	BE410894
40	1034.5	71.4	897	12	BF785457
41	1031.5	71.2	671	10	BE388047
42	1027	70.9	780	13	B1463709
43	1022.5	70.6	809	10	BE384608
44	1017	70.2	635	13	BG974146
45	1011.5	69.8	597	10	BE408222

ALIGNMENTS

RESULT 1
LOCUS BQ212057
DEFINITION BQ212057 899 bp mRNA linear EST 02-MAY-2002
5', mRNA sequence.
ACCESSION BQ212057
VERSION BQ212057.1 GI:20391921
KEYWORDS
SOURCE human.
ORANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabps-r@mail.nih.gov
 Tissue Procurement: ATCC/DCMP/DFP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1331 row: a column: 01
 High quality sequence stop: 675.

FEATURES
 source
 Location/Qualifiers
 1. 899
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6061320"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: PCMV-SPORT6; Site.1: NotI;
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 190 a 234 c 259 g 214 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 2 32e-165 Length: 899
 Score: 1436.00 Matches: 267
 Percent Similarity: 99.26% Conservative: 1
 Best Local Similarity: 98.89% Mismatches: 2
 Query Match: 99.10% Indels: 0
 DB: 14 Gaps: 0

US-09-905-674-2 (1-270) x BQ212057 (1-899)

QY 1 MethisTyrTyrArgTyrSerAsnAlaLysValSerCysTrrPylLysTyrLeuLeuPhe 20
 |||||
 Db 5 ATGCACTAATATAGTACTCTAAAGCCAGGTCAGCTGCTGATACAGTACTCTTTTC 64
 QY 21 SetTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
 |||||
 Db 65 AGCTACAAATCATCTTCTGGTGGCTGGAGTGTCTTCTCTGGAGTGGGCTGTGGCA 124
 QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrArgMethIleAsp 60
 |||||
 Db 125 TGGAGGAAAAGGCTGCTGTCCGACCTCACCMAAGTACCCGGATGATGAAATCGAC 184
 QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
 |||||
 Db 185 CCGTGTGCTGCTGCTGATGTGGGCGGTGGATGTTACCCCTGGGGTTCGCGCTGC 244
 QY 81 ValGlyAlaLeuArgLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 |||||
 Db 245 GTGGGGGCTCTGGGAGATATCTGCTGTCTCACTTTTCTGTGGCACCATCGTCTC 304
 QY 101 IlePhePheLeuGluLeuAlaValAlaValAlaPheLeuPheGluAspTrpValArg 120
 |||||
 Db 305 ATCTTTCTCTGGAGCTGCTGTGGCCGCTGCGCTTCTCTTCCAGGACTGGGTGAGG 364
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
 |||||
 Db 365 GACCGGTTCCGGAGTCTTTCAGAGACATCAAGTCTACCGGGAGCATATCATCTCG 424
 QY 141 GluAsnIleAspSerLeuGluLysAlaAsnGlnCysGlyAlaIleArgProGlu 160
 |||||
 Db 425 CAAACCTATGACTCCCTTCAGAAAGCTAACAGTGTGGGCGCATATGGCCCTTAA 484
 QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
 |||||
 Db 485 GACTGGAGCTCAGCTCTACTTCAATTGCGACGGTGCAGTACAGCCGAGAGATGTC 544

QY 181 GlyValProPheSerCysCysValProAspProAlaGluLysValAlaAsnThrGlnCys 200
 |||||
 Db 545 GGGGTCCCTTCTCTCTGCTGGTGCAGATCTTCGCAAAAAGTGTGGAACACACACTGT 604
 QY 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
 |||||
 Db 605 GGATATGATGTCCAGATTCACAGTCAAGAGCAAGGATGATGATCACTTCACCAAAAGC 664
 QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
 |||||
 Db 665 TGATCCAGGGGGCTGGAAAGCTGGCTCCCGGGAACATTTCATTTGCTGGCTGCTTTC 724
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
 |||||
 Db 725 ATGCGCATCTCGCTGTTCAGATATTTGGCATCTCTCTCGCAAAAGCGATGATCAGAC 784
 QY 261 IleGluAlaValLysAlaGlyHisHisPhe 270
 |||||
 Db 785 ATCNAAGCAGTGAAGAGCGGCATCAGTTC 814

RESULT 2
 LOCUS BM923308 1104 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6625806 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5758901
 5', mRNA sequence.
 ACCESSION BM923308
 VERSION BM923308.1 GI:19373687
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1104)
 REFERENCE
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabps-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12803 row: h column: 06
 High quality sequence start: 7
 High quality sequence stop: 700.

FEATURES
 source

Location/Qualifiers
 1. 1104
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5758901"
 /clone_lib="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber. (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
 BASE COUNT 218 a 315 c 319 g 251 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1 76e-164 Length: 1104
 Score: 1430.00 Matches: 269
 Percent Similarity: 99.26% Conservative: 0

Best Local Similarity: 99.26% Mismatches: 1
 Query Match: 98.69% Indels: 1
 DB: 14 Gaps: 0
 US-09-905-674-2 (1-270) x BM923308 (1-1104)

QY 1 Methistyrtyrargtyrserasnalaalysvalsercystpttyrlystyreuleuphe 20
 |||||
 DB 72 ATGCACCTATTATAGATACCTAAGCCCAAGGTCACCTGCTGTACAGATACCTCTTTTC 131
 QY 21 SerTyrAsnIleIlePheThrPheuAlaGlyValAlaPheLeuGlyValGlyLeuPhe 40
 |||||
 DB 132 AGCTACACATCATCTTCTGTTGGCTGGAGTGTCTTCTTGGAGTGGGCTGTGGGA 191
 QY 41 TTPSerGlyLysGlyValLeuSerAspLeuThrLysValThrMetHisGlyIleAsp 60
 |||||
 DB 192 TGGAGCGAAAGAGTGTGCTGTCCGACCTCACCAAGATGACCGGATGATGAAATCGAC 251
 QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGly 80
 |||||
 DB 252 CCTGTGTGCTGTGCTGTGATGGGCTGTGATGTTCCACCTGGGCTTCCCGGCTTC 311
 QY 81 ValGlyAlaLeuArgLysLysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 |||||
 DB 312 GTGGGGGCTGTGGGAGAAATCTGCTGCTCACTTTCTGTCGACCAATCGTCTC 371
 QY 101 IlePhePheLeuGlyLeuAlaValAlaValAlaPheLeuPheGlyAsnPTTPValArg 120
 |||||
 DB 372 ATCTTCTCTGTGAGCTGTGCTGTGGCGCTGTGGCTCTCTGTTCAGAGATGGGTGAGG 431
 QY 121 AspArgPheArgGlyPhePheGlySerAsnIleLysSerTyrArgAspAspIleAsp 140
 |||||
 DB 432 GACCGGTTCGGAGATCTTCTGAGACCAATCACTACCGGACGATATCGATCTG 491
 QY 141 GluAsnLeuIleAspSerLeuGlyLysAlaAsnGlyCysGlyValAlaTyrGlyProGly 160
 |||||
 DB 492 CAHAACCTCATGCACTCCTTCAGAAAGCTAACCAAGTGTGTGGCCATATGGCCCTGAA 551
 QY 161 AspTyrAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgLysCys 180
 |||||
 DB 552 GACTGGGACCTCAACGCTCACTTCATTTGACGGGTGCCAGCTACAGCGAGAAAGTGC 611
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlyLysValAlaAsnThrGlyCys 200
 |||||
 DB 612 GGGGTCCCTTCTCTCTGCTGCTGCGCAGATCTGCGCAAAAGTTGTGAACACAGTGT 671
 QY 201 GlyTyrAspValArgIleGlyLeuLysSerLysTyrAspGlySerIlePheThrLysGly 220
 |||||
 DB 672 GGATATATGTCAGGATTCAGCTGAGAGCAAGTGGATGATCCATCTTCACGANAAGC 731
 QY 221 CysIleGlyAlaLeuGlySerTyrPheProArgAsnIleTyrIleValAlaGlyValPhe 240
 |||||
 DB 732 TGCATTCAGAGGGCTGGAAAGCTGGCTCCCGCAACATTTACATTTGGCTGCGCTTC 791
 QY 241 IleAlaIleSerLeuLeuGlyIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
 |||||
 DB 792 ATCCCAATCTGCTGTGTCAGATATTTGGCATCTTCTGCGAAGGACGCTGATCTCAGAC 851
 QY 261 IleGlu-AlaValAlaLysAlaGlyHisHisPhe 270
 |||||
 DB 852 ATCGAAGGCACTGAAGCCGCGCATCTTC 882

RESULT 3
 BG674208 866 bp mRNA linear EST 01-MAY-2001
 LOCUS 602619868F1 NCL_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4745512 5',
 DEFINITION mRNA sequence.
 ACCESSION BG674208
 VERSION BG674208.1 GI:13905604
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 866)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
 Plate: LIML0592 row: 9 column: 17
 High quality sequence stop: 821.
 Location/Qualifiers
 1..866
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4745512"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCL_CGAP Library."
 BASE COUNT 176 a 229 c 255 g 206 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.02e-159 Length: 866
 Score: 1383.00 Matches: 262
 Percent Similarity: 98.87% Conserved: 0
 Best Local Similarity: 98.87% Mismatches: 2
 Query Match: 95.45% Indels: 1
 DB: 12 Gaps: 0

US-09-905-674-2 (1-270) x BG674208 (1-866)

QY 7 SerAsnAlaLysValSerCysTyrPtyrLysTyrLeuPheSerTyrAsnIleIlePhe 26
 |||||
 DB 2 TCTAACGCCAAGTCAGCTGCTGTACAAAGTACCTCTTTACAGTACACATCATCTTC 61
 QY 27 TTPLeuAlaGlyValAlaPheLeuGlyValGlyLeuTTPAlaTTPSerGlyLysGlyVal 46
 |||||
 DB 62 TGGTTGGCTGGAGTGTCTCTCTTGTGAGTCCGGCTGTGGGATGGAGCGAAACGGGTGTG 121
 QY 47 LeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProValValLeuValLeu 66
 |||||
 DB 122 CTGTCCGACCTCACCAAGATGACCCGGATGCTATGAAATCGACCTGTGTGCTGCTG 181
 QY 67 MetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGly 86
 |||||
 DB 182 ATGTGGGGGTGTGATGTTACCTGGGGCTCGCCGCGCTGCTGGGGAG 241
 QY 87 AsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlyLeu 106
 |||||
 DB 242 AATATCTGCTTCTCAACTTTTCTGTGGCACATGCTCATCTTCTCTGAGACTG 301
 QY 107 AlaValAlaValLeuAlaPheLeuPheGlyAsnPTTPValArgAspArgPheArgLysPhe 126
 |||||
 DB 302 GCTGTGGCGCTGTGCTCTTCTGTTCCAGGACGTGGTGAAGGACCGGTTCGGGAGTTTC 361
 QY 127 PheGlySerAsnIleLysSerTyrArgAspAspIleAspLeuGlyAsnLeuIleAspSer 146
 |||||
 DB 362 TTGAGAGCAACATCAAGCTTACCGGACGATATCGATCTCAAAACCTCATTCACATCC 421
 QY 147 LeuGlyLysAlaAsnGlyCysGlyValAlaTyrGlyProGlyLysTyrAspLeuAsnVal 166
 |||||
 DB 422 CTTCAGAAAGCTAACAGTGTGTGGCATATGGCCCTGAAGACCTGGGACCTCAACCTC 481
 QY 167 TyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCysGlyValAlaProPheSerCys 186

```

|||||
Db 482 TACTTCATTTGACGCGGTGACCTACAGCCGAGACAGTGGGGGTCCCTTCTCTCC 541
|||
Qy 187 CysValProAspProAlaGlnValValAsnThrGlnCysGlyTyrAspValArgIle 206
|||
Db 542 TGGGTGCGAATCTCTCGCAAAAAGTTGTAACACACACTGTGGATATATGTCAGGATT 601
|||
Qy 207 GlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeuGlu 226
|||
Db 602 CACCTGACAGGCAAGTGGGATGATGTCATCTTCACGAAAGGCTGCATCCAGCGCTGGAA 661
|||
Qy 227 SerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeu 246
|||
Db 662 AGCTGGCTCCCGGCAACATTTACATTTGCGCTGGCTCTTCATCGCATCTCGCTGTTG 721
|||
Qy 247 GlnIlePheGlyIlePheLeuAlaArg-ThrIleLysSerAspIleGlnAlaValLysAl 266
|||
Db 722 CAGATTTTGGCATCTCTTCTGGCAAGGAGCGCTGATCTCAGACATCGAGCGAGTGAAGC 781
|||
Qy 266 aclyhshshspshe 270
|||
Db 782 CGAGCATCACTTC 794
|||
RESULT 4
BQ067138 1125 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_6767382 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5751493
ACCESSION BQ067138
VERSION BQ067138
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1125)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM12784 row: C column: 14
High quality sequence stop: 587.
Location/Qualifiers
1. 1125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5751493"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCWV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 239 a 333 c 297 g 256 t
ORIGIN
Alignment Scores: 1.86e-151 Length: 1125
Pred. No.: 1324.00 Matches: 260
Score: 5

```

```

Percent Similarity: 95.65% Conservative: 4
Best Local Similarity: 94.20% Mismatches: 6
Query Match: 91.37% Indels: 0
DB: 14 Gaps: 6
US-09-905-674-2 (1-270) x BQ067138 (1-1125)
Qy 1 MethslyrTyrArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuLeu 20
|||
Db 93 ATGACATATATAGATACTCTTAACGCCAAGTCAAGTGTGTGACAGTACAGTACCTCTTC 152
|||
Qy 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValIleTrpAla 40
|||
Db 153 AGCTACAAATCATCTTCTGTGTGGTGGAGTGTCTTCTTGAGAGTGGGGCTGGCA 212
|||
Qy 41 TrpSerGlnLysGlyValLeuSerAspLeuThrLysValThrArgMetLysGlyLeasp 60
|||
Db 213 TGGAGCGAAAGGGGTGTGTGCTCGACCTCACCAGATGATGGATGAC 272
|||
Qy 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
|||
Db 273 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 332
|||
Qy 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
|||
Db 333 GTGGGGGCTCTGCGGGAATATCTGCTTCTCACTTTTGTGTGGACCATGTGTCTC 392
|||
Qy 101 IlePhePheLeuGlnLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120
|||
Db 393 ATCTTCTTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 452
|||
Qy 121 AspArgPheArgGluPhePheGlnSerAsnIleLysSerTyrArgAspIleAspLeu 140
|||
Db 453 GACGGGTTCGGGAGTCTTCTGAGACCAATCAATCACTTACCGGAGACATATCATCTTG 512
|||
Qy 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
|||
Db 513 CAAACCTCATGCATCCCTTCAGAAAGCAACACATGCTGTGTGCGCATATGCGCTGAA 572
|||
Qy 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgLysCys 180
|||
Db 573 GACTGGGACCTCAACCTCACTTCAATTCACGCGGTGCCAGTACAGCGAAGAAGTGC 632
|||
Qy 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
|||
Db 633 GGGGTCCCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
|||
Qy 201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
|||
Db 693 GATATGATGTCTCAAGATTCACCTGAGAGACCAAGTGGATGATCCATCTTCAAAAAGGC 752
|||
Qy 221 CysIle-GlnAlaLeuGlnSerTrpLeuProArg-AsnIleTyrIleVal-AlaGlyVal 239
|||
Db 753 TGGATTCACGGCGCTGGAAAGCTGTGCTCCGGGCAACTTTACATTTGTGGCGGGCGTC 812
|||
Qy 240 PheIleAlaIleSerLeu-LeuGlnIlePheGlyIlePhe-LeuAlaArgThrLeuLys 259
|||
Db 813 TTCATGGCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 872
|||
Qy 259 erAspIleGlnAla-ValLysAlaGlyhshspshe 270
|||
Db 873 CAACATCCAGGAGACATAAAGCGCGCATTCATTC 908
|||
RESULT 5
BQ753541 883 bp mRNA linear EST 15-MAY-2001
LOCUS 602733110P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876370 5',
DEFINITION mRNA sequence.
ACCESSION BQ753541
VERSION BQ753541.1 GI:14064194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 883)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabps-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1759 row: d column: 03
 High quality sequence stop: 786.

FEATURES

source
 1..883
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4876370"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."*

BASE COUNT 174 a 238 c 264 g 207 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.61e-149 Length: 883
 Score: 1304.00 Matches: 262
 Percent Similarity: 95.27% Conservative: 0
 Best Local Similarity: 95.27% Mismatches: 8
 Query Match: 89.99% Indels: 7
 DB: 12 Gaps: 0

US-09-905-674-2 (1-270) x BG753541 (1-883)

QY 1 MethISTYRTYRGTYSerASnAlaYSValSerCysTTPYRlySlyLeuLeuPhe 20
 DB 51 ATGCACATATATAGTACTCTACGCCAAGGTCAGCTCTGGTACAACTCCTTTTC 110
 QY 21 SerTYRAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
 DB 111 AGCTACACATCATCTCTCGTTGGCTGAGAGTGTCTTCCTTGAGAGTGGGCTGGGCA 170
 QY 41 TTPSerGlyLysGlyValLeuSerAspIleuThrLysValIThrArgMetHisGlyIleAsp 60
 DB 171 TGGACCGAAAGGCTGCTGTCCGACCTCACCAAGGACCCGATCATGAAATGCAC 230
 QY 61 ProValIleValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyLys 80
 DB 231 CCTGTGGTCTGCTCTCATAGTGGGCTGTGATGTTCACCTGGGGGTTCGGCGCTGC 290
 QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPheCysGlyThrIleValLeu 100
 DB 291 GTGGGGGCTGTGGCGGAATATCTGCTCAACTTTTCTGTGGCAGCATGCTGCTC 350
 QY 101 IlePhePheLeuGlyLeuAlaValAlaValAlaPheLeuPheGlnAspTrpValArg 120
 DB 351 ATCTTCTTCTGGAGCTGTGGCTGTGGCTGTGGCTTCCTTCAGGACTGGGTGAGG 410
 QY 121 AspArgPheArgGlyPhePheGlySerAsnIleLysSerTYRArgAspAspIleAspLeu 140
 DB 411 GACCGGTTCGGGAGGTTCTTCGAGAGCAACATCAAGTCTACCGGAGCATATGATCTG 470

QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrglyProGlu 160
 DB 471 CAAAACCTCATCAGTACTCCCTTCAGAAAGCTAACAGTCTGTGGCGCATATGGCCCTGAA 530
 QY 161 AspTrpAspLeu-AsnValIlePheAsnCysSerGlyAlaSerIleAspArgGlyLysCys 180
 DB 531 GACTGGAGCTCAAGAGTCTACTTCATTAATTCAGAGGCTCAGCTACAGCCAGAGAGTG 590
 QY 180 sGlyValProPheSerCysValPro-AspProAlaGlnLysValAlaAsnThrGln 200
 DB 591 CGGGGTCCCTTCCTCTCTGCTGCGCCAGATCTCTGCGCAAAAGT-CTGACACACAGT 649
 QY 200 ySGlyTYRAspValArgIleGlnLeuLysSerlySTPAspGlySerIlePheThrLys- 219
 DB 650 GTGATATGATGATCAGGATTCAGCTGAGACCAAGTGGAGAGTCCATCTTCAGAAACA 709
 QY 220 GlyCysIleGlnAlaLeuGlu-SerTrp-LeuProArgAsnIleTyrlleValAlaGly 239
 DB 710 GGCTCATCCAGGCGCTGGAAAGCTGCGCTCCCGGAGACATTACATTGTGGCTGGCG 769
 QY 239 alPheIleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuLys 259
 DB 770 TCTTCATCGCATCTCGCTGTGGCCGATATGGCATCTTCTGCGAGGCGCTGATCT 829
 QY 259 eraSPILeGluAlaValLysAlaGlyHisHisPhe 270
 DB 830 CAGACATCGAGGAGTGAAGCCGG--CATCACTTC 862

RESULT 6 857 bp mRNA linear EST 04-OCT-2001
 LOCUS B1835459
 DEFINITION 603087572F1 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5226856 5',
 mRNA sequence.
 ACCESSION B1835459
 VERSION B1835459.1 GI:15947009
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 857)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabps-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1750 row: o column: 17
 High quality sequence stop: 840.

FEATURES

source

1..857
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5226856"
 /clone_lib="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. Insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC library."*

BASE COUNT 159 a 236 c 246 g 216 t


```
|||||
Db 351 ATCTTCTTCCTGAGCTGCTGTGCGCCGTGCGCTTCCTGCTTCCAGAGACTGGGTGAG 410
Oy 121 ASPATGPhaArgGluPhePheGluSerAsnIleYsSerTyrArAspAspIleAspIle 140
Db 411 GACCGGTTCGGAGAGTCTTCGAGAGCACATCAAGTCTTACCGGAGATATCGATCTG 470
Oy 141 GlnAsnIleIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
Db 471 CAAACCTCATGACCTCCCTTCAGAAAGCTAACCAAGTGTGTGGGCGCATATGCCCCTGA 530
Oy 161 AspTTPAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
Db 531 GACTGGACCTCAAGCTCACTTCAATTGACAGCGGTGCGACAGTACAGCGGAGAAAGTGC 590
Oy 181 GlyValProPheSerCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
Db 591 GGGGTCCCTCTCTCTCGCTGCGCAGATCCTGCGCAAAAGTTGTGAACACACAGTGT 650
Oy 201 GlyTyrAspValArgIleGlnLeuLysSerLysTyr-AspGluSerIlePheThrLysGly 220
Db 651 GGATATGATGTCAGATTCAGCTGAAGACAGTGGGAGTGTCCATCTTCACCAAAAG 710
Oy 220 yCys-IleGln-AlaLeuGlnSerTyrLeu-ProArgAsnIleTyrIleValAlaGlyVa 239
Db 711 CTGCATTCCAAGCGCTGGAAGCTGCTCCCGCGGAACATTTACATTTGCTGCGCGT 770
Oy 239 IPhelIleAlaIleSerLeuLeuGlnIle-PheGlyIlePheLeuAlaArgThrLeuIle 259
Db 771 CTTCATTCGGCATCTCGCGTGTTCAGATATTTGGGCAATTCCTCCCTGGAAGGCGCTGATCT 830
Oy 259 er 259
Db 831 CC 832
RESULT 8
BI663945 831 bp mRNA linear EST 12-SEP-2001
LOCUS 60328609P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5322531 5',
DEFINITION mRNA sequence.
ACCESSION BI663945
VERSION BI663945.1 GI:15578178
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS 1 (bases 1 to 831)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1818 row: b column: 04
High quality sequence stop: 789.
FEATURES
source
location/Qualifiers
1..831
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5322531"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
```

```
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 160 a 218 c 239 g 214 t
ORIGIN
Alignment Scores:
Pred. No.: 8,97e-146 Length: 831
Score: 1276.00 Matches: 246
Percent Similarity: 96.15% Conservative: 4
Best Local Similarity: 94.62% Mismatches: 9
Query Match: 88.06% Indels: 3
DB: Gaps: 0
US-09-905-674-2 (1-270) x BI663945 (1-831)
Oy 1 MethIstYrYrArGlyYrSerAsnAlaLysValSerCysTyrIlySyrLeuLeuPhe 20
Db 52 ATGCACTATTTATGATACTCGAACGCCAGGTCAAGTGTGTACAACTGCTCTTC 111
Oy 21 SerTyrAsnIleIlePheThrPheAlaGlyValAlaPheLeuGlyValGlyLeuThrPala 40
Db 112 AGCTACAAATATCGTCTTTGGCTGGCTGGAGTTGTCTTCTTGAGAGTGGGCTGGGCA 171
Oy 41 TrpSerGluYrGlyValAlleuSerAspLeuThrLysValThrArgmeThIsgIlyIleasp 60
Db 172 TGGAGCGAAAAGGTTGTGCTGTCCGACCTCACCAAGTGAACCCGGTTCATCGAATCGAC 231
Oy 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
Db 232 CCCGTGCTGCTGCTGTGATGTGGTGGTGTGATGTTCACACTGGGATTCGCAAGCGCTGT 291
Oy 81 ValGlyAlaLeuArGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
Db 292 GTGGGGGCCCTCCAGAGAACATCGCTGCTCAAGTTTCTGTGGGGCCATTGTGCTC 351
Oy 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTyrValArg 120
Db 352 ATCTTCTTCTCGGAACCTGCGCGTGGCCGCTTTTATTTCCAAAGCTGGGTAGA 411
Oy 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
Db 412 GACCGGTTCGGGAATTCCTTCAGAGCAACATCAAGTCTATCGGATGCATCGACCTG 471
Oy 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
Db 472 CAGAACCTCATTTGACTCCCTTCAGAAAGCTATATCAGTGTGCGGGGCTTACGGCCCTGAA 531
Oy 161 AspTTPAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
Db 532 GACTGGACCTCATGTCATCTTCACTGCAGTGTGCTCCAGCTACACCGAAGAAATGT 591
Oy 181 GlyValProPheSerCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
Db 592 GGGGACCTCTCTCTGCTGTGTCAGATCTCCACAAAAAGCTGGAAACACACACTGT 651
Oy 201 GlyTyrAspValArgIleGlnLeuLysSerTyrTrpAspGluSerIlePheThrLysGly 220
Db 652 GGCTATGATGTCCGATTCACAGTCAAGAGCGAG-TGGGATGATGTCATCTTACAAAAGGA 710
Oy 221 CysIleGlnAlaLeuGluSerTyrPheProArgAsnIleTyr-IleValAlaGlyValPh 240
Db 711 TGCAATCCAGGCTGTGGAAGGCTGCTGCCAGAAACATCTACCAATGTGTGCTGTCTT 770
Oy 240 eIleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSer 259
Db 771 CATGCGCATCTCACTGTGCAAGATTTTGGCATCTTCG-GGGAGGACCTGTGATCTCG 827
RESULT 9
BI663945 742 bp mRNA linear EST 19-JUN-2002
LOCUS 60572112
DEFINITION UT-M-FD0-byd-n-13-0-UT.r1 NIH_BMAP_FD0 Mus musculus cDNA clone
```


cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com>

BASE COUNT 192 a 246 c 277 g 235 t 1 others

ALIGNMENT SCORES:

Pred. No.:	4.57e-145	Length:	951
Score:	1271.00	Matches:	238
Percent Similarity:	99.58%	Conservative:	0
Best Local Similarity:	99.58%	Mismatches:	1
Query Match:	87.72%	Indels:	1
DB:	9	Gaps:	0

US-09-905-674-2 (1-270) x AL519790 (1-951)

```

QY 1 MethISTYTYTArGTYSerASnAlaLysValSerCysTTPYrLYsTYrLeuLeuPhe 20
    |||||||
Db 218 ATGCACCTATTATAGTACTTACCAAGCAAGTCACTGCTGTACAACTCCTTTTC 277

QY 21 SerTYrASnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
    |||||||
Db 278 ACCTACACATCATCTCTCGGTGGCTGAGTGTCTTCTTCTTGAGAGTGGGCTGTGGGCA 337

QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLYsValThrArgMetHisGlyLeuAsp 60
    |||||||
Db 338 TGGAGCGAAAGGTGTGCTGTCCGACCTCACCAAGTGACCCGATGATGATGAAATCGAC 397

QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
    |||||||
Db 398 CCTGTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 457

QY 81 ValGlyAlaLeuArgGlyLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
    |||||||
Db 458 GTGGGGGCTCTGGGGAGAAATCTGTCTGTCACTTTCTGTGGGCAATCGTGCTC 517

QY 101 IlePhePheLeuGlyLeuAlaValAlaValIleAlaPheLeuPheGlnAspTrpValArg 120
    |||||||
Db 518 ATCTTCTTCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 577

QY 121 AspArgPheArgGlyIlePhePheGlyIleSerAsnIleLYsSerTYrArGAspAspIleAspLeu 140
    |||||||
Db 578 GACCGGTTCGGGAGTTCTTCGAGACCAACATCAAGTCTTACCGGAGCATTCGATCTG 637

QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyValAlaTrpGlyProGln 160
    |||||||
Db 638 CAAACCTCATTCGACCTCCCTCGAAGAAAGCTAACCACTGTGTGGCCATATGGCCCTGAA 697

QY 161 AspTrpAspLeuAsnValLYrPheAsnCysSerGlyAlaSerTYrSerArgGlyLysCys 180
    |||||||
Db 698 GACTGGGACCTCAACGCTCACTCAATTCAGCGGTGCGACGTACAGCCAGAGAGAGTGC 757

QY 181 GlyValProPheSerCysValProAspProAlaGlyLysValValAlaAsnThrGlnCys 200
    |||||||
Db 758 GGGGT-CCCTTCTCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 816

QY 201 GlyTYrAspValArgIleGlnLeuLYsSerLYsTrpAspGlyIlePheThrLYsGly 220
    |||||||
Db 817 GGATATGATGTCAAGATTCAAGTGAAGAGCAAGTGGATGATGATTCATCTTCACGAAGGC 876

QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTrpIleValAlaGlyVal 239
    |||||||
Db 877 TGCATCCAGGCGCTGGAAGCTGTCTCCCGGCAACATTTCATTTGTGGTGGCTG 933

RESULT 11
BG913227 797 bp mRNA linear EST 05-JUN-2001
LOCUS 602811894F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4945782
DEFINITION 5', mRNA sequence.

```

```

ACCESSION BG913227
VERSION BG913227.1 GI:14293703
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 797)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/,
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10887 row: 1 column: 23
High quality sequence stop: 685.
Location/Qualifiers
FEATURES
source
1..797
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4943782"
/clone_1lb="NCL_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: brain; Vector: pCMV-SPORT6; Site_1: NCL;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT 160 a 211 c 232 g 194 t
ORIGIN
Alignment Scores:
Pred. No.: 3.51e-142 Length: 797
Score: 1246.50 Matches: 253
Percent Similarity: 95.17% Conservative: 3
Best Local Similarity: 94.05% Mismatches: 4
Query Match: 86.02% Indels: 9
DB: 13 Gaps: 1
US-09-905-674-2 (1-270) x BG913227 (1-797)
QY 1 MethISTYTYTArGTYSerASnAlaLysValSer-CysTTPYrLYsTYrLeuLeuPhe 20
    |||||||
Db 7 ATGCACCTATTATAGTACTTACCAAGCAAGTCACTGCTGTACAACTCCTTTTC 66

QY 20 heserTYrASnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
    |||||||
Db 67 TCAGCTACAACTCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126

QY 40 IatrpSerGlyLysGlyValLeuSerAspLeuThrLYsValThrArgMetHisGlyIleAla 60
    |||||||
Db 127 CATGAGAGGAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 186

QY 60 sPpValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
    |||||||
Db 187 ACCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 246

QY 80 ySValGlyAlaLeuArgGlyLysAsnIleCysLeuLeuAsnPhePheCysGlyThrIleVal 100
    |||||||
Db 247 GCGTGGGGCTTGGGGAGAAATATCTCTGTCTCAACTTTTCTGTGTGTGTGTGTGTGTGTGT 306

QY 100 euIlePhePheLeuGlnLeuAlaValAlaValIleAlaPheLeuPheGlnAspTrpVal 120
    |||||||
Db 307 TCATCTTCTTCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 366

QY 120 rGAspArgPheArgGlyIlePhePheGlyIleSerAsnIleLYsSerTYrArGAspAspIleAsp 140
    |||||||

```

```
|||||
Db 367 GGGGCGGTTCCGGAGTCTTCGAGAGCAACATCAAGTCTACCGGGAGATATCGATC 426
Qy 140 euglnasnleuileaspserrleuglnlysalaasngincysysgylvalatyrglyprog 160
Db 427 TCGAAAACCTCATCGACTCCCTTCAGAAAGCTAACCAAGTGTGTGGCGCATATGGCCCTG 486
Qy 160 lualsprrpaspleuasnvaltyr-pheasnrcysserglyalasertryserargglulysc 180
Db 487 AAGCTGAGGACCTCAACAGTCTACTCAATGACAGCGGTGCACACTACAGCGGAGACAAG 546
Qy 180 ysglyvalpropheser-cyscysvalproasproalaglnlysalvalasnthglnc 200
Db 547 GCGGGGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
Qy 200 ysgly-Tyr-AspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThr 219
Db 607 GTGGCATATCATCATCTCAGAGTTAGCTGAAAGACAGTGGATGACTCCATCTTCACGAA 666
Qy 219 sglycysIleGlnAlaLeuGlu-SerTrpLeu-ProArgAsnIleTyrIle-ValAlaG 238
Db 667 AGCTCATCTCCAGCGCTGGAACAGCTGGCTCCCGCGAACAATTACATGGAGCTGG 726
Qy 238 yValPheIleAlaIleSerLeuLeu---GlnIlePheGlyIlePheLeuAlaArgThr-L 257
Db 727 CGTCTTCATCGCATCTCGCTGGTGGTCGAAATATTTGGGATCTTCTGCAAGAGCGCC 786
Qy 257 euileaserasp 260
Db 787 TGATCTCAGAC 797
RESULT 12
BQ948800
LOCUS BQ948800 870 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT.8806962 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6338432
ACCESSION BQ948800
VERSION BQ948800.1 GI:22364278
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 870)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bhs-rt@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LILCM2531 row: k column: 09
High quality sequence start: 3
High quality sequence stop: 530.
Location/Qualifiers
1..870
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6338432"
/clone_id="NIH_MGC_47"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOT8; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
```

```
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."
BASE COUNT 175 a 223 c 257 g 214 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,04e-142 Length: 870
Score: 1246.50 Matches: 241
Percent Similarity: 94.51% Conservative: 0
Best Local Similarity: 94.51% Mismatches: 9
Query Match: 86.02% Indels: 5
Db: 14 Gaps: 2
US-09-905-674-2 (1-270) x BQ948800 (1-870)
Qy 3 TyrTrpArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuLeuPheSerTyr 22
Db 4 TATTATAGATACCTCAAGCCAGGTCAAGTGTGATCAAGTACCTCTTTCAGCTAC 63
Qy 23 AsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValIleLeuTPrpAlaTrpSer 42
Db 64 AACATCATCTTCTGGTGGCTGGAGTGTCTTCTTGGAGTCGGGCTGGGCAATGGAGC 123
Qy 43 GluysgIyValLeuSerAspLeuThrLysValThrArgMetHisgIyIleAspProVal 62
Db 124 GAAAGGGGTGTGCTGTGCGACCTACCAAACTGACCCGGATGCTATGCAATCGACCTGTG 183
Qy 63 ValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGly 82
Db 184 GTGCTGTCTCGATGATGGTGGCGGTGTGATGTTCACCTGGGGGTTCGCGGCTGGTGGGG 243
Qy 83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIlePhe 102
Db 244 GCTCTGGGGAGATATCTGTGCTCAACTTTTCTGTGGCAACATCGCTCATCTTC 303
Qy 103 PheLeuGlnLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAspArg 122
Db 304 TTCTGTGAGGTGGCTGTGGCGGTGCTGCTCTCTGTCCAGGACATGGTGAGGACCGG 363
Qy 123 PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn 142
Db 364 TTCGGGAGTCTTCGAGAGCAACATCACTACCGGAGCAATATCATCTGCAAAAC 423
Qy 143 LeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProgIuAspTrp 162
Db 424 CTGATTCAGATCCCTTCGAAAGACTAACCAAGTGTGTGGCCATATGGCCCTGAAGACTGG 483
Qy 163 AspLeuAsnValTyrPheAsnrcysSerGlyAlaSerTyrSerArgGluLysCysGlyVal 182
Db 484 GACCTCAACGCTCTACTTCATTTGAGCGGTGCCAGCTACAGCGGAGAAAGTGGGGGTC 543
Qy 183 ProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyr 202
Db 544 CCTTCTCTGCTGCTGCGGCGAGATCTCGCAAAAGTTGTGAACACACAGTGTGATAT 603
Qy 203 AspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIle 222
Db 604 GATGTCAAGATTCAGCTGGAAGACCAAGTGGATGATCTTCACGAAAGGCTCATC 663
Qy 223 GlnAlaLeuGlnSerTrp-LeuProArg-AsnIleTyrIleValAla---GlyValPhe 241
Db 664 CAGCGCGTGGAAAGCTGGGCTCCCGGGAACATTTACATTTGGGGTGGCGGTCTCTCA 723
Qy 241 leuIleaser-----LeuLeuGlnIlePheGlyIlePhe 252
Db 724 TTGCGCATCTCCGCTGTGTCANAAATATTTTGGCATCTTT 764
RESULT 13
BQ765719 755 bp mRNA linear EST 15-MAY-2001
BQ765719
LOCUS 602739738F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4869633 5'
DEFINITION mRNA sequence.
```

ACCESSION	BG765719
VERSION	BG765719.1
KEYWORDS	EST.
SOURCE	GI:14076372
ORGANISM	Homo sapiens human. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health. Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: cgaqbs-remail.nih.gov
TITLE	Tissue Procurement: ATCC/DCTD/DPF
JOURNAL	cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMLN at: http://image.lnl.gov
COMMENT	Plate: LMCML741 row: k column: 10 High quality sequence stop: 754.
FEATURES	location/qualifiers 1..755
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4869633" /clone_id="NIH_MGC_49" /tissue_type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(S). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." 1"
BASE COUNT	143 a 207 c 226 g 179 t
ORIGIN	Library. "1"
Alignment Scores:	
Pred. No.:	6.53e-142 Length: 755
Score:	1244.00 Matches: 228
Percent Similarity:	100.00% Conservatlve: 1
Best local Similarity:	99.56% Mismatches: 0
Query Match:	85.85% Indels: 0
DB:	12 gaps: 0
US-09-905-674-2 (1-270) x BG765719 (1-755)	
OY	1 MethistyrtyrArGyrsSerAsnalalysValserCystripTylyLeulephe 20
Db	68 ATGCACATATTATATAACTCTAAAGCCAAAGGTGAGTGCTGGTCAAGTACTCTTTTC 127
OY	21 SerTyraSnilelelpehrpleualaglYvalValheuleuglyValglyLeutrpAla 40
Db	128 AGCTACACATCATCTCTTGtggctgGcgAggtttCTTCTTGAGTCGGGCTGGGCCA 187
OY	41 TrpSerGlulysglYalleSerAspdeThryValThrtraqMeThisglYIlasp 60
Db	188 TGGAGCGAAAAGSGTGTGCTGCCGACCTCACCAAAGTAGAACCGGATGATGAATGAC 247
OY	61 ProValValleuVallleuMeValglYvalValmethehrleuglyPhealaglyCys 80
Db	248 CCGTGTGTGCTGGCTCGAATGTGGGGGTGTGTATGTTCAACCCTGGGGTTCCGCGGCTGC 307
OY	81 ValglYalaleuArGluasnilleCyseLeuleuasnaphPhecysglYthrllleValleu 100
Db	308 GTGGGGGCTCTGGGGGAGAAATATCTGCTTCTCAACTTTTCTGTGACACCATCTGCTC 367
OY	101 IlePhePheleugtlueualavalAlaValleuAlaPheleuPheclInasPrtpValArg 120

QY	Db	Accession	Version	Keywords	Source	Organism	Location/Qualifiers
QY 368	ATCTCTTCCTCGAGCGCTGCTGTGGCCGAGCTGGCTTCTCTGTTCCAGGACTGAGG	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 121	AsparagphearaglunherheglusersanllylsSerTyrArgaspallleaspleu	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 428	GACCGGTCCTCGAGGATTCCTCGAGCAACATCAATCAATCTACCGGACATATTCATCTG	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 141	GlunsluenuleakpserleuglnulysalaasnglnCysCysglyAlaTyrArglyProglu	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 488	CAAAACCTATGACATCCCTTCAGAAAGCTAACCACTGCTGTGGCAGATATG6CCCTGGA	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 161	AspTTPAspleuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArglnulysCys	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 548	GACTGGAGCCTCAACAGCTTCACTTCAATTCAGAGGGTCCAGCTACAGCCGAGGAAGTGC	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 181	GlYValProPhSerSerCysValProAspProAlaGlnLysValAlaAsnThrGlnCys	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 608	GGGGTCCCTCTCTCCGTCGCGTGCAGTCTCTGCGCAAAAGTGTGTAAACACAGTGT	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 201	GlYTyRAspValAlaGlnLeuLysSerLysTTPAspGlnSerLephethrlysgly	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 668	GGATATGATGTCAGAGTTCAGCTGGAAGCACTGGAGTGCATCTTCACGAAAGT	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 221	CyslleGlnAlaLeuGlnSerTTPleu	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 728	TGCATCCAGGCGCTGGAAGCTGTGTC	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	RESULT 14	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	LOCUS	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	DEFINITION	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	ACCESSION	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	VERSION	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	KEYWORDS	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	SOURCE	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	ORGANISM	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	REFERENCE	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	AUTHORS	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	TITLE	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	JOURNAL	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	COMMENT	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	FEATURES	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',
QY 14	SOURCE	BG760633	602717387/1	NIH_MGC_49	Homo sapiens	CDNA clone	IMAGE:4840916 5',

Alignment Scores:

Pređ. No.: 6.72e-140

Score: 1228.00

Percent Similarity: 97.54%

Best Local Similarity: 97.548

Query Match: 84.75%

DB: 12

Length: 810

Matches: 238

Conservative: 0

Mismatches: 6

Indels: 2

Gaps: 0

US-09-905-674-2 (1-270) x BG760633 (1-810)

QY	1	MethistyrTrpArgTrpSerAsnAlaLysValSerCysTrpTrpLysTyrLeuLeuPhe	20
Db	68	ATGCACATTTATAGTAACTCTTACGCCAAGGTCAAGTCTGTGTGTAACACTACCTCTTTTC	127
QY	21	SerTrpAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla	40
Db	128	AGCTACAAACATCATCTCTCGTTGGCTGGAGTGTCTTCTTGAGTGGCGGTGGGCA	187
QY	41	TrpSerGluLysGlyValLeuSerAspLeuThrLysValIleTrpArgMetHisGlyIleAsp	60
Db	188	TGAGACGAAAAGGGTGTGCTGTCCACCTCCACCAAGGACCCGGATTCACAGAAATCGAC	247
QY	61	ProValIleValLeuValLeuMetValGlyValAlaIleMetPheThrLeuGlyPheAlaGlyCys	80
Db	248	CCTGTGTGTCTGTCTCTGATGTGGGCGGTGGTGAATGTTTACCTCGGGGTTGGCCGGCTGC	307
QY	81	ValGlyAlaLeuAlaArgLysAsnIleCysLeuLeuAsnPhePheCysGlyIleValLeu	100
Db	308	GTGGGGGCTCTCGGGAGAAATATCTGCTGTGCTCACTTTTCTGTGGCACCATCATCTGCTC	367
QY	101	IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg	120
Db	368	ATCTCTCTCTGTGAGCTGTGCTGTGGCCGTGGCTTCTGTCTGTTCAGAGACTGGGTGAGG	427
QY	121	AspArgPheArgGluPhePheGluSerAsnIleLysSerTrpArgAspAspIleAspLeu	140
Db	428	GAACGGTTCGGGAGATTCTTCAGAGCAACATCAAGTCTTACCGGAGCATATCATCTG	487
QY	141	GlnAsnIleuIleAspSerLeuGluLysAlaAsnGlnCysGlyValAlaTyrGlyProGlu	160
Db	488	CAAAACCTCATGACATCCCTTCAGAAAGCTAACCAAGTGTGGGCGCATATGNGCTGAA	547
QY	161	AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTrpSerArgGluLysCys	180
Db	548	GACTGGGACTCAACTCTACTTCAATTGCACGGTGCAGCTTAAGCCGAGAGAAAGGC	607
QY	181	GlyValProPheSerCysValProAspProAlaGlnLysValAlaAsnThrGlnCys	200
Db	608	GGGGTCCCTTTCCTGCGTGCAGTACTCTGGCGCAAAAAGTTGTGACACACAGTGT	667
QY	201	GlyTrpAspValArgIleIleuLysSerLysTrpAspGluSerIlePheThrLysGly	220
Db	668	GGATATGATGTACAGATTAGCTGGAAGACCAAGTGGATGATCTTTCACGAAAGGG	727
QY	221	CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe	240
Db	728	TG-ATTCAGGCGGTGGAAAGGCTGGGTCCGGCGGAACATTACATTGTGTGT-GGCGCTTC	785
QY	241	IleAlaIleSer 244	
Db	786	ATCGCATCTCG 797	
RESULT 15			
LOCUS	BC974843	770 bp	mRNA linear EST 12-TUN-2001
DEFINITION	602842725F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978446 5',		
ACCESSION	BC974843		
VERSION	BC974843.1	GI:14362480	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota, Metazoa: Chordata; Craniata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Mus I (bases 1 to 770)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D.

found through the I.M.A.G.E. Consortium/LLNL at:
 DNA Sequencing by: Incyte Genomics, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 CLONAL LIBRARY PREPARATION: TATE GENOMICS/LLNL

FEATURES

source

```

/organism="Mus musculus"
/strain="NMR1"
/db_xref="taxon:10090"
/clone="IMAGE:4978446"
/clone_lib="NCI_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennigshausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

```

BASE COUNT
ORIGIN

ORIGIN

Alignment Scores:

Pred. No.:	2.55e-139
------------	-----------

Score: 1223.00

Percent Similarity: 97.06%

Best Local Similarity: 95.80%

Query Match: 84.40%

DB: 13

Length: 770

Matches: 228

Conservative: 3

Mismatches: 5

Indels: 2

Gaps: 0

US-09-905-674-2 (1-270) x BG974843 (1-770)

QY	1	MeHhISyTYrTYrATyTSerASnaAlaYsAlSerCySTrPYrLYsTYrLeuPhe	20
Db	53	ATGCATATTATAGATTACTCGAACGCCGAGGTCAGACTCTGGTCAAGTACTCTCTTC	112
QY	21	SeTYrAsnIleIlePheTrPLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrAla	40
Db	113	AGCTCAAAATATGCTCTTTGGCTGGCTGAGATGTCCTTCCTTGAGTGGGCTGGGCA	173
QY	41	TrpSerGluLYsGlyValLeuSerAspLeuThrLYsValAlrArGMeThIeGlyIleAsp	60
Db	173	TGGAGCGAAACGGGTGTGTGCTGTCCACCTCACCAAGGTGACCCGGTTCATGCAATGCAC	233
QY	61	ProValValLeuValLeuMetValGlyValAlaMeCPheTrPheLeuGlyPheAlaGlyCys	80
Db	233	CCCGTGTCTGTCTTGATGTGTGGCGGTGGTGGATGTTCAACATGGCATTTGGCAGCGCT	293
QY	81	ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu	100
Db	293	GTCGGGGCCCCCTCGAGAGACATCTGCCTCGCTCAAGTTTTTCTGTGGGGCATTTGTCTC	353
QY	101	IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg	120
Db	353	ATCTTCTTCTGTGAACTGGCCGCGTGGCGGTTGGCCTTTTATTATCCAGAGCTGGGTGAGA	413
QY	121	AspArgPheArgGluPhePheGluSerAsnIleLYsSerTYrArgAspAspTLeuAspLeu	140


```

Db      413 GACCGGTTCCGGGAATCTTCGAGACAACATCAAGTCCTATCGGGATGACATCGACCCTG 472
QY      141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
         |||
Db      473 CAGAACTCATTTGACTCCCTTCAGAAAGCTAATCAGTGTGGGGCTTACGGCCCTGAA 532
QY      161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
         |||
Db      533 GACTGGAGACCTCAATGTCTACTTCAACTGCAAGTGTGGCCAGCTACAGCCGAGAGAAATGT 592
QY      181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
         |||
Db      593 GGGGTACCCCTTCTCTGCTGTGTGCGCAGATCCTGCACAAAAGTGTGACACACACAGTGT 652
QY      201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGly 220
         |||
Db      653 GGCTATGATGTCCGATTCAGCTGAAGAGCAGAGTGGGATGAGTTCATCTTTACAAAAGGA 712
QY      221 CysIleGln-AlaLeuGluSerTrpLeuProArg-AsnIleTyrIleVal 236
         |||
Db      713 TGCATCCAAAGGCTCTGGAAGGCTGGCTGCCGCAACATCTACATTGTG 762

```

Search completed: December 19, 2002, 07:45:49
 Job time : 2012 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 19, 2002, 06:19:07 ; Search time 51 Seconds

(without alignments)
1623.583 Million cell updates/sec

Title: US-09-905-674-2

Perfect score: 1449

Sequence: 1 MHYRYSNAMVSCWYKYLFL.....IFLARLISDIKAVKAGHHF 270

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Xgapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+lp2n.model -DEV=xlh
-o=/cgn2_1/USPTO/spool/US0905674/runat_16122002.111846.6791/app_query.fasta.1.455
-DB=Issued.patents.NA -QEMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USER=US0905674.ecgn.1.1.17.6runat_16122002.111846.6791 -NCPU=6 -ICPU=3
-NO_XLPUX -NO_MAP -LARGEXERT -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

Issued.patents.NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	278	19.2	1151	2	US-08-807-044-2
2	266	18.4	827	4	US-09-333-599-5
3	266	18.4	870	4	US-09-333-599-1
4	264.5	18.3	1344	3	US-08-705-771-8
5	259.5	17.9	1452	2	US-08-807-044-4
6	259.5	17.9	1452	5	PCR-US91-04986-1
7	257.5	17.8	687	1	US-08-254-493-2
8	257.5	17.8	687	1	US-08-408-222B-2
9	257.5	17.8	1120	1	US-08-408-222B-3
10	257.5	17.8	1624	4	US-08-430-225A-19
11	252.5	17.4	977	2	US-08-855-140-2
12	252.5	17.4	1120	1	US-08-254-493-3

13	218.5	15.1	933	3	US-08-808-148-2	Sequence 2, Appl1
14	218.5	15.1	1289	4	US-09-020-956-111	Sequence 111, App
15	218.5	15.1	1289	4	US-09-030-607-111	Sequence 111, App
16	218.5	15.1	1289	4	US-09-605-785-111	Sequence 111, App
17	218.5	15.1	1289	4	US-09-439-313-111	Sequence 111, App
18	218.5	15.1	1289	4	US-09-352-616A-111	Sequence 111, App
19	218.5	15.1	1289	4	US-09-232-149A-111	Sequence 111, App
20	198.5	13.7	1001	3	US-08-705-771-6	Sequence 6, Appl1
21	197.5	13.6	399	3	US-09-188-930-62	Sequence 62, Appl
22	197.5	13.6	399	3	US-09-188-930-243	Sequence 243, App
23	176.5	12.2	801	4	US-09-020-956-16	Sequence 16, Appl
24	176.5	12.2	801	4	US-09-030-607-16	Sequence 16, Appl
25	176.5	12.2	801	4	US-09-605-785-16	Sequence 16, Appl
26	176.5	12.2	801	4	US-09-439-313-16	Sequence 16, Appl
27	176.5	12.2	801	4	US-09-352-616A-16	Sequence 16, Appl
28	176.5	12.2	801	4	US-09-232-149A-16	Sequence 16, Appl
29	164	11.3	398	4	US-09-385-982-273	Sequence 273, App
30	161.5	11.1	560	4	US-09-221-298-12	Sequence 12, Appl
31	152.5	10.5	740	4	US-09-020-956-17	Sequence 17, Appl
32	152.5	10.5	740	4	US-09-030-607-17	Sequence 17, Appl
33	152.5	10.5	740	4	US-09-605-785-17	Sequence 17, Appl
34	152.5	10.5	740	4	US-09-439-313-17	Sequence 17, Appl
35	152.5	10.5	740	4	US-09-352-616A-17	Sequence 17, Appl
36	152.5	10.5	740	4	US-09-232-149A-17	Sequence 17, Appl
37	147.5	10.2	1966	4	US-08-957-130-14	Sequence 14, Appl
38	141.5	9.8	1782	4	US-09-149-476-120	Sequence 120, Appl
39	136.5	9.4	751	4	US-09-020-956-12	Sequence 12, Appl
40	136.5	9.4	751	4	US-09-030-607-12	Sequence 12, Appl
41	136.5	9.4	751	4	US-09-605-785-12	Sequence 12, Appl
42	136.5	9.4	751	4	US-09-439-313-12	Sequence 12, Appl
43	136.5	9.4	751	4	US-09-352-616A-12	Sequence 12, Appl
44	136.5	9.4	751	4	US-09-232-149A-12	Sequence 12, Appl
45	133	9.2	729	4	US-09-020-956-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-807-044-2
; Sequence 2, Application US/08807044
; Patent No. 5863735
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN TRANSMEMBRANE 4 SUPERFAMILY
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTED for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807, 044
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0224 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

QY	159	ProguIsapTTPrAspLeuAsnValTtYrPhaSnCysserolYalaserTYrSerArgJlu	178
Db	537	TCACAGGACCTGGCGAGACAGAGTGGATCCGC-----TCACAGGAG	578
QY	179	Lyscsgly-----ValProPhSeScyScyValProAsProAlaGlnLysValVal	196
Db	579	GCCGGTGGCCCTGTGGTCGCCAGACAGCTGCTGC-----AAGACGGTG	620
QY	197	AsnThrGlnCySgly-----TyrAspValArgIleGlnLeu	208
Db	621	GTGGCTCTTTGTGGACAGCAGACCATGCTCCAAATCTACAAAGTGGAG-----	671
QY	209	LysSerLysTTPrAspGlnSerIlePhetThrLysJLysIleGlnAlaLeuGlnSerTTP	228
Db	672	-----GGCGGCTGCATCCACCAAGTTGGAGACCTTC	701
QY	229	LeuProArGAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuGlnIle	248
Db	702	ATTCAGGAGACCTGAGAGGTGATTTGGGGCTGTGGGATCGGCACTGGCTGTGCAGAGT	761
QY	249	PheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlnAlaValLysAlaGlnHis	268
Db	762	TTTGGCATGATCTTCAGCTGCTGCTG-----TACAGAGGTCTCAACGCTGAGAC	812
QY	269	His	269
Db	813	TAC	815
RESULT 3			
US-09-333-599-1			
Sequence 1, Application US/09333599			
Patent No. 6245898			
GENERAL INFORMATION:			
APPLICANT: Testa, Jacqueline E.			
APPLICANT: Quigley, James P.			
APPLICANT: Seandel, Marco			
TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS			
TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS			
FILE REFERENCE: SUNY			
CURRENT APPLICATION NUMBER: US/09/333,599			
CURRENT FILING DATE: 1999-06-15			
NUMBER OF SEQ ID NOS: 5			
SOFTWARE: Patentln Ver. 2.0			
SEQ ID NO 1			
LENGTH: 870			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-333-599-1			
Alignment Scores:			
Pred. No.:	6,9e-24	Length:	870
Score:	266.00	Matches:	74
Percent Similarity:	44.48%	Conservative:	51
Best Local Similarity:	26.33%	Mismatches:	108
Query Match:	18.36%	Indels:	48
DB:	4	Gaps:	11
US-09-905-674-2 (1-270) x US-09-333-599-1 (1-870)			
QY	6	TySerAsnAlaLysValSerCyS-----TTPrLysTyrLeuLeuPheSerTyr	22
Db	110	TTTCAAGAGGAGAGACACACATATGGCCACCCGTTTGCTCAAGTACACCGTGTATTACTAC	169
QY	23	AsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValAlaGlyLeuTrpAlaTrpSer	42
Db	170	AATTCGCTCTTCTGGCTGGCGCTGGCTGCTGCTATGGCACTGGGACATCTGGACGCTGGCC	229
QY	43	GluLysGly---ValLeuSerAspIleThrLysValThrArgMetHisGlyIleAspPro	61
Db	230	CTCAAGAGGTGACTACATCAGCGCTGCTGGCTCAAGGACACTAC-----CTGGGCACA	280
QY	62	ValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCySVal	81

```

Db      261   :..... ||| ||||: ||| |||||
OY      82   GlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhenPheCysGlyThrIleValLeu 101
Db      341   GCCACCTTCAGAGAGCGCGGAACTCTGTGGCCCTGATCACTTCAATCCGCCTCATCATC 400
OY     102   PhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAsp 121
Db      401   TTTTGCTGGGAGATCATCGCTGTAATCTCGCTCACCCCTACTACCAAGCAGTGAACACG 460
OY     122   ArgPheArgGluPhePheGlnSerAsnIle---LysSerTyrrArgsp-----AspIle 138
Db      461   GAGCTCAAGGAGAACCCTGAAGACACATGACCAAGCGCTACACACAGCCGGGCCATGAG 520
OY     139   AspleuGlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnLysCysGlyAlaTyrgly 158
Db      521   GCTGTGACACACGCGCTGTGGACACCTGCAGAGAGATTCCACTGCTGTGGCAGACAACAC 580
OY     159   ProGluAspTrpAspPheLeuAsnValTyrrPheAsnCysSerGlyAlaSerTyrrSerArgGlu 178
Db      581   TCACAGGACTGCGCAGACAGTAGAGTGGATCCGC-----TCACAGGAG 622
OY     179   LysCysGly-----ValProPheSerCysCysValProAspProAlaGlnLysValAl 196
Db      623   GCCGCTGGCCCTGTGGTCCAGACAGCTGCTGC-----AAGACGGTc 664
OY     197   AsnThrGlnCysGly-----TyrrAspValArgIleGlnLeu 208
Db      665   GTGGCTCTTTGTGGACAGCGAGACCAATGCCCTCCACATCTACCAAGTGGAG----- 715
OY     209   LysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeuGlnUsterTrp 228
Db      716   -----GGCGGCTGCATCACCAAGTTGGAGACCTTC 745
OY     229   LeuProArgAsnIleTyrrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIle 248
Db      746   ATCCAGGAGCACCTGAGGGCTATTGGGAGCTGTGGGATGGCAATGCTGTGTGCAGGT 805
OY     249   PheGlyIlePheLeuAlaArgThreLeuIleSerAspIleGluAlaValLysAlaGlyHis 268
Db      806   TTGGCATGATCTTCAGCTGTGCTG-----TACAGAGATCTCAAGCTGAGAAC 856
OY     269   His 269
        ::
Db      857   TAC 859

RESULT 4
US-08-705-771-8
; Sequence 8, Application US/08705771
; Patent No. 6054289
; GENERAL INFORMATION:
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
; APPLICANT: Jian Ni and Jing-Shan Hu
; TITLE OF INVENTION: Human Genes, Sequences and
; TITLE OF INVENTION: Expression Products
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,771
; FILING DATE: August 30, 1996

```


Db 203 -----TTCCAAACCTCCCTCCCTCAGCGTGGGCAATGTTGTCTCATCTGGGCTCT 256
QY 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90
Db 257 ATTATATGTAAGTTCCTCCCTCCGCGCTCATGAGGCTCTATCAAGAAACAAGTGTCTG 316
QY 91 LeuAsnPhePheCysGlyThrIleValIleuIlePhePheLeuGluLeuAlaValAla 110
Db 317 CTATATGTCGTTCTTCATCCCTGCGTGTGATATTCCTCTGAGAGGTGACCTGGCCATC 376
QY 111 LeuAlaPheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsn 130
Db 377 CTGCTCTTTGATATGAACAAGAAAGGATATGCTGCTAAGGCTGTGACCGACAGC 436
QY 131 IleYsSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAla 150
Db 437 ATCCACCGTTACCACTCAGACAATAGCACCAAGCAGCGCTGGAGACTCCATCCAGTCAATT 496
QY 151 AsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCys 170
Db 497 CTGCAGTGTGTGTGATAAATGCGACAGATGTGG----- 532
QY 171 SerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysCysValProAsp 190
Db 533 -----ACCACTGGCCCAACGACATCTTGC----- 556
QY 191 ProAlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSer 210
Db 557 CCCTCAGATCGAAGAGT----- 574
QY 211 LysTrpAspGluSerIlePheThrIysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230
Db 575 -----GAGGTGCTATGCAAGCAAGCAAGACTGTGGTTTCA 610
QY 231 ArgAsnIleTyrIleValAlaGlyAlaPheIleAlaIleSerLeuLeuGlnIlePheGly 250
Db 611 TCCAAATTCCTGTATATGCGAATCATCAACACCATCTGTATGTGTGATGAGGTGTGGGG 670
QY 251 IlePheLeuAlaArgThrLeuIleSerAspIleGlu 262
Db 671 ATGTCTTGTGCACTGACCTCGAATCGCCAGATTGAC 706

RESULT 6
PCT-US91-04986-1
Sequence 1, Application PC/TUS9104986
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Allen, Janet
APPLICANT: Aruffo, Alejandro
APPLICANT: Camerini, David
APPLICANT: Lauffer, Leander
APPLICANT: Oquendo, Carmen
APPLICANT: Simmons, David L.
APPLICANT: Stamenkovic, Ivan
APPLICANT: Stengel, Siegfried
APPLICANT: Amiot, Martine
TITLE OF INVENTION: Rapid Immunoselection Cloning Method
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee & Associates
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/04986
FILING DATE: 19910715

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/553,759
FILING DATE: 13-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/498,809
FILING DATE: 23-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/379,076
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/160,416
FILING DATE: 25-FEB-1988
ATTORNEY/AGENT INFORMATION:
NAME: Wall, Margaret M.
REGISTRATION NUMBER: 33,462
REFERENCE/DOCKET NUMBER: 11-88C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1452 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 74..733
PCT-US91-04986-1

Alignment Scores:
Seq. No.: 9.83e-23 Length: 1452
Score: 259.50 Matches: 62
Percent Similarity: 42.86% Conservative: 46
Best Local Similarity: 24.60% Mismatches: 89
Query Match: 17.91% Indels: 55
DB: 5 Gaps: 6

US-09-905-674-2 (1-270) x PCT-US91-04986-1 (1-1452)

QY 16 LysTrpLeuLeuPheSerTyrAsnIleIlePheThrLeuAlaGlyValAlaPheLeuGly 35
Db 101 AAGATATGCCCTGTTTCTTCAACTGCTCTTTGGAGTCTGTGGCTCTGCAATTTTGGC 160
QY 36 ValGlyLeuTrp-----AlaTrpSerGluLysGlyValLeuSerAspLeuThrLysVal 53
Db 161 TTTGGGATCTACCTGCTGATCCACACAACTTCGAGTGTCTC----- 202
QY 54 ThrArgMetHisGlyIleAspProValValLeu-----ValLeuMetValGlyVal 70
Db 203 -----TTCCATTAACCTCCCTCCCTCAGCGTGGGCAATGTTGTTCATCTGGGGCTCT 256
QY 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90
Db 257 ATTATATGTAAGTTCCTCCCTCCGCGCTCATGAGGCTCTATCAAGAAACAAGTGTCTG 316
QY 91 LeuAsnPhePheCysGlyThrIleValIleuIlePhePheLeuGluLeuAlaValAla 110
Db 317 CTATATGTCGTTCTTCATCCCTGCGTGTGATATTCCTCTGAGAGGTGACCTGGCCATC 376
QY 111 LeuAlaPheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsn 130
Db 377 CTGCTCTTTGATATGAACAAGAAAGGATATGCTGCTAAGGCTGTGACCGACAGC 436
QY 131 IleYsSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAla 150
Db 437 ATCCACCGTTACCACTCAGACAATAGCACCAAGCAGCGCTGGAGACTCCATCCAGTCAATT 496
QY 151 AsnGlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCys 170
Db 497 CTGCAGTGTGTGTGATAAATGCGACAGATGTGG----- 532

RESULT 8

US-08-408-222B-2

Sequence 2, Application US/08408222B

Patent No. 5776727

GENERAL INFORMATION:

APPLICANT: Ikeyama, Shuichi

APPLICANT: Koyama, Masaru

APPLICANT: Miyake, Masayuki

APPLICANT: Senoo, Masaharu

TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dike, Bronstein, Roberts & Cushman

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/408,222B

FILING DATE: 22-MAR-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/254,493

FILING DATE: 06-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-079996-1991

FILING DATE: 12-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-085396-1991

FILING DATE: 14-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP-022321-1992

FILING DATE: 07-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 41777-DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 687 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to genomic RNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

ORGANISM: human

CELL TYPE: breast carcinoma

CELL LINE: ZR-75-1

US-08-408-222B-2

Alignment Scores:

Pred. No.:	5,58e-23	length:	687
Score:	257,50	Matches:	64
Percent Similarity:	43.72%	Conservative:	44
Best Local Similarity:	25.91%	Mismatches:	94
Query Match:	17.77%	Indels:	45
DB:	1	Gaps:	6

US-09-905-674-2 (1-270) x US-08-408-222B-2 (1-687)

QY 16 LysTyrIleuLeuPheSerTyrAsnIleIlePheThrPheuAlaGlyValValPheLeuGly 35

```

Db 31 AAATACCTGCTGCTGATTAACCTGATCTTGGCTGCGGAGTGCCTGCTGCGC 90
QY 36 ValGlyLeuThrPalaIatPserGlu-----LysGlyValLeuSerAspLeuThrLys 52
Db 91 ATGGACTATGGCTCCGATTCGACTTCAGACCAAGACATCTTCGACGAAGAACTAAT 150
QY 53 ValThrArgMetHisGlyIleAspProValValLeuValLeuMetValGlyValValMet 72
Db 151 AATAATATTCACCTTCTACACAGAGTCAATATTCATTCGATCGAGCGCGGCGCCATG 210
QY 73 PheThrIleuGlyPheAlaGlyCysValGlyValAlaLeuArgGluAsnIleCysLeuLeuAsn 92
Db 211 ATGCTGGTGGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 270
QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluIleuAlaValAlaValAla 112
Db 271 CTGTTCTTGGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
QY 113 PheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGlnIleuSerAsnIleLys 132
Db 331 TATTCACCAAGATGAGTGTATTAAGAAAGTCCAGAGATTTTACAAAGACACCTACAC 390
QY 133 SerTyrArg---AspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsn 151
Db 391 AAGCTGAACCAAGATGAGTCCAGCGGGAACGCTGAAGACCATCCATGCGCTTG 450
QY 152 GlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSer 171
Db 451 AACTGCTGTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 172 GlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
Db 481 -----CAGTTTATCTCAGACATCTGC-----CCC 504
QY 192 AlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerLys 211
Db 505 AAGAAGCAGCTGCTGCAAC----- 525
QY 212 TrpAspGlnSerIlePheThr---LysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230
Db 526 -----TTCCCGGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
QY 231 ArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGly 250
Db 571 AATAATTCACATCATCGGCGCAGTGGCATTCGCGATTCGCGTGCATGATATTGCG 630
QY 251 IlePheLeuAlaArgThrLeu 257
Db 631 ATGATCTTCAGTATGATCTTG 651

RESULT 9
US-08-408-222B-3
Sequence 3, Application US/08408222B
Patent No. 5776727
GENERAL INFORMATION:
APPLICANT: Ikeyama, Shuichi
APPLICANT: Koyama, Masaru
APPLICANT: Miyake, Masayuki
APPLICANT: Senoo, Masaharu
TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
```

```

SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,222B
FILING DATE: 22-MAR-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,493
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-079996-1991
FILING DATE: 12-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-085396-1991
FILING DATE: 14-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-022321-1992
FILING DATE: 07-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 41777-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1120 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: breast carcinoma
CELL LINE: ZR-75-1
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 112...795
OTHER INFORMATION: E Mat peptide
US-08-408-222B-3

Alignment Scores:
Pred. No.: 1,186-22 Length: 1120
Score: 257.50 Matches: 64
Percent Similarity: 43.72% Conserves: 64
Best Local Similarity: 25.91% Mismatches: 94
Query Match: 17.77% Indels: 45
DB: 1 Gaps: 6

US-09-905-674-2 (1-270) x US-08-408-222B-3 (1-1120)
QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValValPheLeuGly 35
DB 142 AATATACCTGCTTGGATTACTTCTGCTGCGGAGATGCTGCTCC 201
QY 36 ValGlyLeuTrpAlaTrpSerGlu-----LysGlyValIleSerAspLeuThrIys 52
DB 202 ATTGACATATGCGTCGACCTGACCTGACACACAGACATCTTGACCAAGAAACTAT 261
QY 53 ValThrArgMetHisGlyIleAspProValValLeuValLeuMetValGlyValValMet 72
DB 262 AATAATAATTCAGCTTACACAGAGATATATTCGATCGGAGCCGCCCTCATG 321
QY 73 PheThrIleuGlyPheAlaGlyCysValAlaLeuAlaGluAsnIleCysLeuLeuAsn 92
DB 322 ATGCTGGGGGCTGCTGCGGCGGCTGTCGAGGAGATCCCATGCTGCTGGGA 381
QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValLeuAla 112
DB 382 CTGTTCTTCGCGCTTCTGCTGATATTCGCCCATGGAATAAGCTGGGCGCATCTGGGGA 441
```

```

QY 113 PheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGlnSerAsnIleIys 132
DB 442 TATTCGCCAAGATGAGGTATTAAGAGAGTCCAGGAGTTTACAAAGACCTACAAAC 501
QY 133 SerTyrArg---AspAspIleAspLeuGlnAsnIleAspSerLeuGlnLysAlaAsn 151
DB 502 AAGCTGMAAACCAAGATGAGCCCGGCGGAAACGCTGAAAGCCATCCTACTATGCGTGG 561
QY 132 GlnCysCysGlyAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSer 171
DB 562 AACTGCTGTGGTTGGTGGGCGGTGGAA----- 591
QY 172 GlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
DB 532 -----CAGTTTATCTCAGACATCTGC-----CCC 615
QY 192 AlaGlnLysValAlaValAsnThrGlnCysGlyTyrAspValArgIleGlnLeuLysSerIys 211
DB 616 AAGAGAGAGCTACTCGAACC----- 636
QY 212 TrpAspGluSerIlePheThr---LysGlyCysIleGlnAlaLeuGluSerTrpLeuPro 230
DB 637 -----TTACCGCTGAAGTCTGCTGATGCCATCAAGAGAGCTCTCGAC 681
QY 231 ArgAsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGly 250
DB 682 AATAAATTCACATCATCGCGGCGGAGTGGCATGCGGATGCGGATGATATTTGGC 741
QY 251 IlePheLeuAlaArgThrLeu 257
DB 742 ATGATCTCTCAGTATGATCTTG 762

RESULT 10
US-08-430-225A-19
Sequence 19, Application US/08430225A
Patent No. 6204000
GENERAL INFORMATION:
APPLICANT: Dong, Jin-Tang; Barrett,
TITLE OF INVENTION: DIAGNOSTIC METHODS AND
TITLE OF INVENTION: GENE THERAPY USING REAGENTS DERIVED FROM THE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,225A
FILING DATE: 28-APR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```

; TOPOLOGY: linear
US-08-430-225A-19

Alignment Scores:
Pred. No.: 2,086-22      Length: 1624
Score: 257.50           Matches: 67
Percent Similarity: 46.54% Conservative: 54
Best Local Similarity: 25.77% Mismatches: 116
Query Match: 17.77%      Indels: 23
DB: 4                    Gaps: 7

US-09-905-674-2 (1-270) x US-08-430-225A-19 (1-1624)

QY 16 LysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGly 35
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 193 AATACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252

QY 36 ValGlyLeuTrpAlaTrpSerGluLys-----GlyValLeuSerAspLeuThr 51
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 253 TTCGGGTGTGGATCTCTGCGCAAGAGACAGTTCATCTCTCTCTCTCTCTCTCTCTCTCC 312

QY 52 LysValThrArgMetHisGlyIleAspProValIleLeuValLeuMetValGlyValVal 71
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 313 AACTGCTTAGGATG-----GGGGCTATGTCTTCATCTGCGGTGGGGGCACTG 360

QY 72 MetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeu 91
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 ACTATGCTATGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 420

QY 92 AsnPhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValLeu 111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 GGGCTGACTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 480

QY 112 AlaPheLeuPheGluAspTrpValArgAspArgPheArgGluPhePheGluSerAsnIle 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 TTCTACTTCAACATGGGCAAGGTGAAGACAGAGATGGCGGCATCTGACTGAGCTCACT 540

QY 132 LysSerTyr-----ArgAspAspIleAspLeuGluAsnIleAspSerLeuGln 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 CGAGACTACACAGACAGATCGGAGAC---AGCTGACAGAGATGCTGGGACTACGTGAG 597

QY 149 LysAlaIleAsnGlnCysCysGlyAlaTrpGlyProGluAspTrpAspLeuAsnValTrpPhe 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 598 GCTCAGGTGAAGTGTGGGCTGCGTCACTCTTCACTGACAGACAGACGTGAGCTC 657

QY 169 AsnCysSerGlyAlaSerTyrSerArgGluLysCysGlyValProPheSerCysVal 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 658 -----ATGATCGCCCTGAGTCACTTACCTCCTCTCTCTCTCTCTCTCTCTCTCTC 699

QY 189 ProAspProAlaGlnLysValAlaValAsnThrGlnCysGlyTyr-----AspVal 204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 AAGGGGGAAGAGACACACAGCTTCTGTGAGAGAGGCTTCTGAGAGCCCGCCGAC 759

QY 205 ArgIleGlnLeu---LysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGln 223
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 760 AAGACCCGAGAGTGGACACCCCTGAGGACTGCGTCTGACAGAGAGGCTGATGAG 819

QY 224 AlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIle 243
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 820 AAGGTGACAGGCTGCTGACAGAGACCTGGGATCATCTCTGCGGTGGGCTGGGTG 879

QY 244 SerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGluAla 263
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 880 GCCATCATGACCTCTGCGGATGTCTCTGTCATCTGCTGTGCGGCGGACGTCACTTCC 939

RESULT 11
; Sequence 2, Application US/08855140
; Patent No. 5854022
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.

```

```

; APPLICANT: Guejter, Karl J.
; TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/855,140
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0296 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MYCONOT01
; CLONE: 779308
US-08-855-140-2

Alignment Scores:
Pred. No.: 4,046-22      Length: 977
Score: 252.50           Matches: 61
Percent Similarity: 46.70% Conservative: 45
Best Local Similarity: 26.87% Mismatches: 90
Query Match: 17.43%      Indels: 31
DB: 2                    Gaps: 7

US-09-905-674-2 (1-270) x US-08-855-140-2 (1-977)

QY 13 CysTrpTyrLysTyrLeuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValVal 32
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 TGCCTTGAACATGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199

QY 33 PheLeuGlyValGlyLeuTrpAlaTrpSerGluLysGlyValLeuSerAspLeuThrLys 52
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CTGCTGGAGTGGGCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 253

QY 53 ValThrArgMetHisGlyIleAspProValIleLeuValLeuMetValGlyValAlaMet 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 254 TCCCCAGCTTCCCTGTTGTCTGACAGCCACCTGCTGATCCGCAATGAGCACCATTTGTC 313

QY 73 PheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeuAsn 92
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 314 ATGCTGACGGCTTCTCTGCGTCTCTGCGGCGCATCAAGGAAACAAAGTCTCTCTCTC 373

QY 93 PhePheCysGlyThrIleValLeuIlePhePheLeuGluLeuAlaValAlaValAla 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 374 AGCTTTTCAATCGCTGCTGTGCTGACCCCTCTGACAGAGTGAATTTACCATCTCTTTC 433

QY 113 PheLeuPheGluAspTrpValArgAspArgPheArgGluPhePheGluSerAsnIleLys 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db 637 -----TTACACCGTGAAGTCTGTCTGATGTCATCAAGAGTCTTCGAC 681
QY 231 ArgasnileTyrlleValaIaglyValaPhelelelleSerleuenglnlePhegly 250
Db 682 AATRAATTCACATCATCGCCGCAAGTGGCATCGGCGATTCGCGTGTGATGATATTGGC 741
QY 251 IlePheleuAlaArgThrleu 257
Db 742 ATGATCTTCAGTATGATCTTG 762

RESULT 13
US-08-808-148-2
; Sequence 2, Application US/08808148
; Patent No. 6020478
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer
; APPLICANT: Goli, Surya
; TITLE OF INVENTION: NOVEL HUMAN TUMOR-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,148
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0218 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 933 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT26
; CLONE: 2187263
US-08-808-148-2

Alignment Scores:
Pred. No.: 6,83e-18 Length: 933
Score: 218.50 Matches: 72
Percent Similarity: 44.09% Conservative: 51
Best Local Similarity: 25.81% Mismatches: 112
Query Match: 15.08% Indels: 45
DB: 3 Gaps: 8

US-09-905-674-2 (1-270) x US-08-808-148-2 (1-933)
QY 7 SerAsnAlaLysValSerCys-----TrpTyrIlyTyrlleuLeuPheSerTyrAsnIle 24
Db 116 ACAGAGCACCACATCATGCTTCAGCTTCATTAGACCAAGATGATGCTCTTCATATTG 175

```

```

QY 25 IlePheTrpLeuAlaIaglyValaPheleuIglyValIleuTrpAlaTrpSerGlnLys 44
Db 176 CTCATCTTCTGTGTGTGTGTCACCTCTGTGGCAGTGGCATCTGGGTGC-AATCGATGG 234
QY 45 GlyValLeuSerAsp-----LeuThrLysVal-ThrArgMetHisGlyIleAsp 61
Db 235 GGCATCTTCTGGAAGATCTTCGGGCCACTGTGCGCAGTGCATGCAATGCTTTCACAGT 294
QY 61 oValValLeuValLeuMetValGlyValaIaIaIaPheThrleuIglyPheAlaIglySva 81
Db 295 GGGCTACTTCTCCTGAGCGAGCGCGTGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTA 354
QY 81 lGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhethecysGlyThrIleValLeuI 101
Db 355 TGGTCTTAAGACTGTGAGACAGGTGCGCTGTGACGTCTTCTTCATCTCTCCCTCAT 414
QY 101 ePhePheleuIglyValaIaIaIaIaPheleuPheGlnAspTrpValArgas 121
Db 415 CTTCATTTGCTGAGTGTGAGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY 121 pArgPheArgGluPhePheGlnSerAsnIle-----LysSerTyrAr 135
Db 460 CACAATGGCTGAGCACTTCTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
QY 135 gAspAspIleAspLeuGlnAsnLeuIleAspSerleuGlnLysAlaAsnGlnCysGly 155
Db 520 TTCCCGAGAAAGCTTCTCACTCAAGTGTGAGACACACACCAAGGCTCAAGTCTGTGG 579
QY 155 yAlaTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTy 175
Db 580 CTTCACCAACTATACGAGATTGTGAGGACTACCTACTCTC----- 619
QY 175 rSerArgLysCysGlyValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 192
Db 620 -----AAAGAGAACAGTGGCTTTCCTCCCATCTGTGTGAATGACCAACGCTCAACAGC 675
QY 192 aGlnLysValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 212
Db 676 CAATGAACCTGCACACCAAGGCTCAACGACCAAAAGTAGAG----- 721
QY 212 pAspGlnSerIlePheThrLysGlyCysIleGlnAlaIaIaIaIaIaIaIaIaIaIaIa 232
Db 722 -----GGTTCCTCAATCAGCTTTGTGATGACATCGCAACTAA 759
QY 232 nIleTyrlleValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 252
Db 760 TGCAGTCAAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819
QY 252 eLeuAlaArgThrleuIleSerAspIleGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIa 270
Db 820 TGTGTCCATGTATCTGATCTGATCATCTACATAAGTC-----CACTTC 862

RESULT 14
US-09-020-956-111
; Sequence 111, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillin, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121,427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO.: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 1289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
OS-09-020-956-111

```

Alignment Scores:	
Pred. No.:	1,12e-17
Score:	218.50
Percent Similarity:	44.09%
Best Local Similarity:	25.81%
Query Match:	15,068
DB:	4
	Gaps: 6

US-09-905-674-2 (1-270) x US-09-020-956-111 (1-1289)

Oy	7	SerSrnAlaLysValSerCys-----TrrPlyLysTrpLeuNeuphesrTyraNile	24
Db	111	ACAGAGCCACCATGCACTGCGTTCCACTTATTAAAGCACATGATGATCCTTCATATTGG	170
Oy	25	IlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpPheIarTrpSerGluLys	44
Db	171	CTCATCTTTTCGTGTGGTGGAAGCCCTTTTGGCAGTGGGCATCTGGGGTGC--AATGCATGG	229
Oy	45	GlyValLeuSerAsp-----LeuThrLysVal-ThrArgMetHisGlyLysLeuSpr	61
Db	230	GGCATTCTTCTTGAAGATCTTCGGGCCACTGCTCCACAGTGCATGCATGGATTGTCAACGT	289
Oy	61	ovalValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysVa	81
Db	230	GGGCTACTTCTCTATCCGACGCCGGCGGTtRGCTTTTGCTCTTGtRGTTCTCGGGCTGCTA	349
Oy	81	IcGlyAlaLeuArGgluNsniLecYsLeuLeuAsnPhePhcCysGlyThrlleValLeuIl	101
Db	350	TGTGTCTTAAGACTGAGACAAGATGTGCCCTGTCACGTTCTTCTTCATCCTCCTCCAT	409
Oy	101	ePheHeuLeuIueLuAlaValAlaValleuAlaPheLeuPheGlnSprPylArGas	121
Db	410	CTTCACTTGTCAAGGTGGAGCTGCTGTGTGGTCCCTTGtGTAC-----AC	454
Oy	121	pARgPheArGgluPheGlnSerAsnile-----LysSerTyraR	135
Db	455	CACATGTGCTAGGACCTTCCTGACGTTGCTGTGATGGCTGCCATCAAGAAGAATATTGG	514
Oy	135	qASpAsPileaspleuGlnAsnLeuIlleasPserLeuGlnLysAlaasnGlnCysGsl	155
Db	515	TTCCCAGAGAACCTTCATCAAGGTGTGGAAACCCACATCAAGAAAGGCTCAAGTCTGTGG	574
Oy	155	ValaIeTrGlyProGluAsPrPaSPleuAsnValTyTr-PheasnCysSerGlyAlaSerTy	175
Db	575	CTTCAACCAATATACGGATTGTGAGACTCAACCTTACTTC-----	614
Oy	175	rSerArGgluLysCysGlyAlaPropheSerCysCys-----ValProAspProAl	192
Db	615	----AAAGGAACAATGGCTTTCCGCCATCTGTGTGCATGACAAACGCAACCAACAGAC	670
Oy	192	aGlnLysValValAsnThrGlnCysGlyTyTraspValArGlleGlnLeuLysSerLyStr	212

Db	671	CAATGAACCTGCACCAAGCAAAAGGCTCACGACCAAAAGTAGAG-----	716
Qy	212	paspqluserlrephenhrrysgylsileglnalaenqluserlrpreuproargas	232
Db	717	-----GGTCTCTTCATCACCTTTTGATACATCCGAACTAA	754
Qy	232	niletyrllevalalaglyalphelealalleleulerleuenglnelepeglyleph	252
Db	755	TGCAGTCAACCTGGGTGTGTGGCAGCTGCAATGGGGGCTCGACGTGGCTGCACATGAT	814
Qy	252	eleuamlaqrhrleulesesapllleglnalaivallyalaglylshlsiphe	270
Db	815	TGTGTCAATGATCTGTACTGCATCTACACATAAGTC-----CACTTC	857
RESULT	15		
	US-09-030-607-111		
	: Sequence 111, Application US/09030607		

RESULT 15
US-09-030-607-111
: Sequence 111, Application US/09030607

```
?
? GENERAL INFORMATION:
? APPLICANT : Xu, Jiangchun
? APPLICANT : Dillon, Davin C.
? TITLE OF INVENTION : COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
? NUMBER OF SEQUENCES : 224
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SEED and BERRY LLP
? STREET: 6300 Columbia Center, 701 Fifth Avenue
? CITY: Seattle
? STATE: WA
? COUNTRY: USA
?
```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patentin Release #1.0, Version #1.30
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/09/030,607
8  FILING DATE: 25-FEB-1998
9  CLASSIFICATION:

```

```

? TELECOMMUNICATION INFORMATION:
? REFERENCED/LOCATED INFORMATION:
?
? TELEPHONE: (206) 622-4900
?
? TELEFAX: (206) 682-6031
?
? INFORMATION FOR SEQ ID NO: 11:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 1289 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: single
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: cDNA
?
? ORIGINAL SOURCE:
?
? ORGANISM: Homo sapiens
?

```

Alignment Scores:	
Pred. No.:	1.12e-17
Score:	218.50
Percent Similarity:	44.09%
Best Local Similarity:	25.81%
Query Match:	15.08%
DB:	4
	8
Length:	128
Matches:	72
Conservative:	51
Mismatches:	112
Indels:	45
Gaps:	8

US-09-905-674-2 (1-270) x US-09-030-607-111 (1-1289)

[illegible]

Db	171	CTCATCTTCTGTGGTGGAGACCCCTGTTGGCATGGGCAATCTGGGTCT-AATCATGG	229
Qy	45	GLYValLeuSerAsp-----LeuThrLysVal-ThrArgMetHisGlyIleAspPr	61
Db	230	GGCAATCTTTTGTGAAGATCTTGGGGCCACTGCTCCACAGTGCATGCAGATTGTTCACAGT	289
Qy	61	oValValLeuValLeuMetValGLYValValMetPheThrLeuGlyPheAlaGlycysVa	81
Db	290	GGGCTACTTCTCTATCCGACGCCGGCGCTTGTGTGCTTTGCTCTTGTGTTCCTGGGCTGCTA	349
Qy	81	IGLYAlaLeuArgGluAsnIleCysLeuLeuAsnPheCysGlyThrIleValLeuI	101
Db	350	TGGTGCTAAAGCTAGAGACAGAGTGTGCCCTGTGACGCTTCTTCATCTCCTCCAT	409
Qy	101	ePhePheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGlnAspTyrValArgAs	121
Db	410	CTTCAATGCTAGGTGTGACGCTGTGGTGGTGGCTTCCTGGTGTAC-----AC	454
Qy	121	PaRgPheArgGluPheGluSerAsnIle-----LysSerTyrAl	135
Db	455	CACAAATGGCTAGGACATCTCCGACAGTGTGCTGATGCTCCGATCAAGAAACATTAATG	514
Qy	135	gAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGsl	155
Db	515	TTTCCAGAAACACTTCACTCAAGGTGTGAACACCACTCAAGAGGCTCAAGTGTGG	574
Qy	155	yAlaTyrGlyProGluSprPaspLeuAsnValTyrPheAsnCysSerGlyAlaSerTy	175
Db	575	CTTACCAACATATACGGAATTTTGGACCTACCCCTACTTC-----	614
Qy	175	rSerArgGluLysCysGlyAlaProPheSerCys-----ValProAspProAl	192
Db	615	-----AAAGAGAACAGTCCCTTCCCATCTCTGTGAATGACACAGTACCAACACAGC	670
Qy	192	agLysValValAsnThrGlnCysGlyTyrTrpAspValArgIleGlnLeuLysSerLysTr	212
Db	671	CAATGAACCTCGACCAAGCAAAAGGCTCCAGCCAAAAGGTGAG-----	716
Qy	212	pAspGluSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTyrLeuProArgAs	232
Db	717	-----GGTTCCTCAATCACTTGTATAGACATCCGAACATA	754
Qy	232	nIleTyrIleValAlaGlyValPheIleAlaIleSerLeuLeuGlnIlePheGlyLePhe	252
Db	755	TGCAGTCACCGTGGGTGGTGTGGCAGCTGAATTTGGGGGCTGAGCTGAGCTGCATGAT	814
Qy	252	eLeuAlaArgThrLeuLeuSerAspIleGlyAlaValLysAlaGlyHisHisPhe	270
Db	815	TGTGTCAATGATCTGTACTGACATATACATATAAGTC-----CACTTC	857

Search completed: December 19, 2002, 07:46:59
Job time : 56 secs

Db	1289	TTTGGGTCATCCATCTCTGTGGGTGGCCGTGGGTAGAGGAGCCACAGCGCTGGACAGC	1348
Qy	1321	GCACTCTCTCCATCAAGCAAGCAGCATGGGGGCCCTTCCGCTAAACGGAGGCG	1376
Db	1349	GCACTCTCTCTCCATCAAGCAAGCAGCATGGGGG--CCTGGCCGTAAAGCGAGGCG	1402
RESULT 3			
AAZ65259	standard; DNA: 2672 BP.		
AAZ65259			
23-MAR-2000	(first entry)		
Human secreted protein gene 10.			
Human: secreted protein; cancer; tumour; developmental abnormality; foetal deficiency; blood disorder; immune system disorder; inflammation; autoimmune disease; allergy; Alzheimer's disease; cognitive disorder; schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder; atherosclerosis; diabetes; cardiovascular disorder; kidney disorder; digestive disorder; endocrine disorder; infection; AIDS; leukaemia; therapy; chromosome 10; ds.			
Homo sapiens.			
MO9958660-A1.			
18-NOV-1999.			
06-MAY-1999;	99WO-US09847.		
12-MAY-1998;	98US-0085093.		
12-MAY-1998;	98US-0085094.		
12-MAY-1998;	98US-0085105.		
12-MAY-1998;	98US-0085160.		
18-MAY-1998;	98US-0085906.		
18-MAY-1998;	98US-0085920.		
18-MAY-1998;	98US-0085921.		
18-MAY-1998;	98US-0085922.		
18-MAY-1998;	98US-0085923.		
18-MAY-1998;	98US-0085924.		
18-MAY-1998;	98US-0085928.		
18-MAY-1998;	98US-0085925.		
18-MAY-1998;	98US-0085927.		
(HUMA-) HUMAN GENOME SCI INC.			
Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA, Olsen HS, Shl Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW, Endress GA, Ehner R;			
WPI: 2000-062296/05.			
P-PSDB; AA176133.			
New isolated human genes and the secreted polypeptides they encode, useful for diagnosis and treatment of e.g. cancers, neurologic disorders, immune diseases, inflammation or blood disorders			
Claim 1; Page 301-302; 47pp; English.			
AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.			
This sequence was found to be present on human chromosome 10.			
AAZ65124 to AAZ65223 represent the secreted proteins encoded by the 97 human genes. The genes and their corresponding secreted polypeptides are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 97 genes, based on which tissues they are most highly expressed in, and include developing products for the			

[illegible]

|||||
Db 965 CAGTGAAGCCGGCCATCATCTTCTAGAGAGCAGAGTTAGAGGAGCCGAGCTGAGCCAGCG 1024
Oy 914 TGGAGAGCCAGAGACCTTTCTGTCATCAGCCCTGATCAGAGGAGAGAGAGCCGAGAC 973
Db 1025 TGGAGAGCCAGAGACCTTTCTGTCATCAGCCCTGATCAGAGGAGAGAGAGCCGAGAC 1084
Oy 974 CCCCAGAGCCAGTGGCCCATCTTAAAGCATCAGCTGAGTGAACCTCTCTGTTCTGCTTG 1033
Db 1085 CCCCAGAGCCAGTGGCCCATCTTAAAGCATCAGCTGAGTGAACCTCTCTGTTCTGCTTG 1144
Oy 1034 CTGGTCTGAAGAACCAGAGGTGCCCCCTGTTACTGCCCCAAACTTGTGATGATCCCTC 1093
Db 1145 CTGGTCTGAAGAACCAGAGGTGCCCCCTGTTACTGCCCCAAACTTGTGATGATCCCTC 1204
Oy 1094 TGGAGTCTACCCAGAGACAGAGATGTGCTTATGTGGAGTGGTGACTCTGAAGACGA 1153
Db 1205 TGGAGTCTACCCAGAGACAGAGATGTGCTTATGTGGAGTGGTGACTCTGAAGACGA 1264
Oy 1154 GAGAGGGCTCTGTGGCTGCCAGAGGGCTTGACTCAGACCCCTGACGCTCAAGCATGT 1213
Db 1265 GAGAGGGCTCTGTGGCTGCCAGAGGGCTTGACTCAGACCCCTGACGCTCAAGCATGT 1324
Oy 1214 CTGCAAGGACACCCCTGCTGCCCTCTCCATGTGCTGCTCCAGACATCTGCTTGGGTCATCA 1273
Db 1325 CTGCAAGGACACCCCTGCTGCCCTCTCCATGTGCTGCTCCAGACATCTGCTTGGGTCATCA 1383
Oy 1274 CATCTGTGGTNGCCGCTGGGTAGAGGGACCCACAGGGGTGAGACGGGATCTCTCTCA 1333
Db 1384 CATCTGTGGTNGCCGCTGGGTAGAGGGACCCACAGGGGTGAGACGGGATCTCTCTCA 1443
Oy 1334 TCAGCAAAAGCAGCATGGGGGGCTTGGCTTAACGGGAGCG 1376
Db 1444 TCAGCAAAAGCAGCATGGGGGG--CCTGCCCGTAAGGAGGCG 1484

RESULT 4
AAFP0628
ID AAFP0628 standard; cDNA: 2715 BP.
XX
AC AAFP0628;
XX
DT 04-MAY-2001 (first entry)
XX
DE Human TANGO 339 cDNA, SEQ ID NO:1.
XX
KW Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW activity modulator; expression modulator; cancer; immunological disorder;
KW cytosolic; immunomodulatory; gene therapy; ss.
OS Homo sapiens.
XX
PN WO200109162-A2.
XX
PD 08-FEB-2001.
XX
PF 31-JUL-2000; 2000WO-US20935.
XX
PR 30-JUL-1999; 99US-0365164.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PI Fraser CC, Sharp JD, Kirst SJ, Barnes TW, Wrighton N, Myers PS;
PI Pan Y;
XX
DR WPI: 2001-138647/14.
XX
P-PSDB: AAB87034, AAB87035, AAB87036.
XX
PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
XX disorders -
PS Claim 2; Page 219-220; 332pp; English.

XX
CC The invention relates to novel secreted/transmembrane proteins, and
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC TANGO 353, TANGO 358, TANGO 365, TANGO 366, TANGO 369, TANGO 383, TANGO
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC murine TANGO 393 is also included within the scope of the invention. The
CC invention also encompasses fragments and variants of the proteins of the
CC invention, and nucleic acids encoding them. The invention additionally
CC relates to host cells comprising a nucleic acid of the invention; methods
CC for the production of a protein of the invention; an antibody specific
CC for a protein of the invention; methods for detecting a protein or
CC nucleic acid of the invention; and methods of identifying agents which
CC bind to or modulate the activity of a protein of the invention. The novel
CC secreted proteins, nucleic acids encoding them, and antibodies against
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression or activity of the secreted
CC proteins. The secreted proteins of the invention may also be used to
CC identify modulators of expression or activity, which may be useful in
CC the treatment of disorders associated with the proteins of the
CC invention e.g., cancers and immunological disorders. The present
XX sequence represents cDNA encoding human TANGO protein of the invention.
SQ Sequence 2715 BP; 559 A; 760 C; 715 G; 681 T; 0 other;

Query Match 93.1%; Score 1292.2; DB 22; Length 2715;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1317; Conservative 2; Mismatches 6; Indels 3; Gaps 2;

Oy 49 GATTCTGCTTCTCAGAAAGATGACATATTATAGATCTTAACGCCAAAGTACAGTCTG 108
Db 192 GATCTCTCTCTCAGAAAGATGACATATTATAGATCTTAACGCCAAAGTACAGTCTG 251
Oy 109 TACAAGTACCTCCCTTTACACTACACATCATCTCTGCTGGCTGAGATGCTCTCT 168
Db 252 TACAAGTACCTCCCTTTACACTACACATCATCTCTGCTGGCTGAGATGCTCTCT 311
Oy 169 GGAATCGGGCTGTGGGATGAGCGAAAGGATGCTGCCACCTCAACAAAGTGACC 228
Db 312 GGAATCGGGCTGTGGGATGAGCGAAAGGATGCTGCCACCTCAACAAAGTGACC 371
Oy 229 CGGATGATGGAATCGACCCCTGTGTGCTGCTCTGATGGGCTGTGATGTTACCC 288
Db 372 CGGATGATGGAATCGACCCCTGTGTGCTGCTCTGATGGGCTGTGATGTTACCC 431
Oy 289 CTGGGGTTCGCCGCGCCGCGGGGGGCTCTCGGAGAAATATCTGCTTGTCAACTTTTC 348
Db 432 CTGGGGTTCGCCGCGCGCGGGGGGCTCTCGGAGAAATATCTGCTTGTCAACTTTTC 491
Oy 349 TGTGGCACATCGTGTCTCTCTCTGAGACTGAGTGGCGCTGCTGAGCTTTC 408
Db 492 TGTGGCACATCGTGTCTCTCTCTGAGACTGAGTGGCGCTGCTGAGCTTTC 551
Oy 409 TTCCAGGACTGGGTGAGGAGCCGTTCCGGAGTTCTTGAGAGCAACATCAAGTCTTAC 468
Db 552 TTCCAGGACTGGGTGAGGAGCCGTTCCGGAGTTCTTGAGAGCAACATCAAGTCTTAC 611
Oy 469 CGGGAGCATATGATCTGCAAAACCTCATGAGTCTGCTCAGAAAGCTTAACAGTCTCT 528
Db 612 CGGGAGCATATGATCTGCAAAACCTCATGAGTCTGCTCAGAAAGCTTAACAGTCTCT 671
Oy 529 GGGCATATGGCCCTGAAGACTGGACCTCAAGTCTACTTCAATTTGACGCGGTGGCAGC 588
Db 672 GGGCATATGGCCCTGAAGACTGGACCTCAAGTCTACTTCAATTTGACGCGGTGGCAGC 731
Oy 589 TACAGCCGAGAGAAGTGCGGGGTCCCTTCTCTGCTGCTGCGAGATCCTGGCGAAAA 648
Db 732 TACAGCCGAGAGAAGTGCGGGGTCCCTTCTCTGCTGCTGCGAGATCCTGGCGAAAA 791
Oy 649 GTTGTGAACACAGTGTGATATGATGACAGATTCAGTGAAGGCAAGTGGATGAG 708
Db 792 GTTGTGAACACAGTGTGATATGATGACAGATTCAGTGAAGGCAAGTGGATGAG 851
Oy 709 TCCATCTTCAAGAAAGCTGCATCCAGGCGCTGGAAGAGTGGCTCCCGCGGAACATTAC 768

```
DB 852 TCCATCTTCAGAAAGGCTGCATCCAGGCGCTGGAAAGCTGCTCCCGGCAACATTAC 911
OY 769 ATTTGGCTGGCGCTTTCATGCCATCTCGCTGTTGCAGATATTGGCATCTTCTGCA 828
DB 912 ATTTGGCTGGCGCTTTCATGCCATCTCGCTGTTGCAGATATTGGCATCTTCTGCA 971
OY 829 AGGATGCTGATCTCGACATGAGGACATGTAAGGCGGCGCTTCATCTTCTGAGAGAG 888
DB 972 AGGAGCTGATCTCGACATGAGGACATGTAAGGCGGCGCTTCATCTTCTGAGAGAG 1031
OY 889 TTGAGGAGCGGAGCTGAGCCAGCGCTGGAGGCCAGAGCCCTTCTCTCCATCAGCCCTA 948
DB 1032 TTGAGGAGCGGAGCTGAGCCAGCGCTGGAGGCCAGAGCCCTTCTCTCCATCAGCCCTA 1091
OY 949 GGTCCAGAGGAGAGAGACCCAGACACCCCAAGAGCAGTGGCCCATCTTAAGCATCAGCGT 1008
DB 1092 GGTCCAGAGGAGAGAGACCCAGACACCCCAAGAGCAGTGGCCCATCTTAAGCATCAGCGT 1151
OY 1009 GACGTGACCTCTCTGTTCTGCTGCTGGTCTGTAAGACCAAGGAGTCCCTTGTACT 1068
DB 1132 GACGTGACCTCTCTGTTCTGCTGCTGGTCTGTAAGACCAAGGAGTCCCTTGTACT 1211
OY 1069 GCCCAACTTGTACTGATCCCTCTGAGTCTACCCAGAGACAGAGATGTCTTTAT 1128
DB 1212 GCCCAACTTGTACTGATCCCTCTGAGTCTACCCAGAGACAGAGATGTCTTTAT 1271
OY 1129 GTGGAGGTGTACTCTAAGAGACAGAGAGGCTCTGCTGCTGCGAGGAGGCTTGACT 1188
DB 1272 GTGGAGGTGTACTCTAAGAGACAGAGAGGCTCTGCTGCTGCGAGGAGGCTTGACT 1331
OY 1189 CAGACCCCTGAGCTCAAGCATGTCTGCAAGACACCCCTGCTCCCTCTCAATGGGCT 1248
DB 1332 CAGACCCCTGAGCTCAAGCATGTCTGCAAGACACCCCTGCTCCCTCTCAATGGGCT 1390
OY 1249 CCAGACATCTGCTTGGGTCAATCACAATCTGTTGGGTGCGGTGAGAGGACCCACA 1308
DB 1391 CCAGACATCTGCTTGGGTCAATCACAATCTGTTGGGTGCGGTGAGAGGACCCACA 1450
OY 1309 GGGGTGAGAGGCGCATCTCTCCATCAAGACAGACATGGGGGCGCTTGCGGTAAAC 1368
DB 1451 GGGGTGAGAGGCGCATCTCTCCATCAAGACAGACATGGGGGCGCTTGCGGTAAAC 1508
OY 1369 GGGAGGCG 1376
DB 1509 GGGAGGCG 1516

RESULT 5
AAS92016
ID AAS92016 standard; cDNA; 2623 BP.
XX
AC AAS92016;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27820.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
```

```
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR MPI, 2001-639362/73.
XX
DR P-PDB; ABG27829.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 27820; 103bp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AB64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 2623 BP; 515 A; 713 C; 719 G; 659 T; 17 other:
XX
Query Match 91.9%; Score 1275.4; DB 23; Length 2623;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1345; Conservative 2; Mismatches 40; Indels 8; Gaps 4;
OY 1 CTTCCTGGCGGAGCGGCGCGCGCTGCTCCCGCGCGCGCGGATTCGTTCT 60
DB 1 CTTCCTGGCGGAGCGGCGCGCGCTGCTCCCGCGCGCGGATTCGTTCT 60
OY 61 CAGAGATGCACTATTATGATATCTAAGCCCAAGTACAGTGTGTATCAAGTACTTC 120
DB 61 CAGAGATGCGCTATGATATCTAAGCCCAAGTACAGTGTGTATCAAGTACTTC 120
OY 121 CTTTTCAGTACAACATCATCTCTGTTGGCTGAGATTGCTTCTTGGAGTGGGCTG 180
DB 121 CTTTTCAGTACAACATCATCTCTGTTGGCTGAGATTGCTTCTTGGAGTGGGCTG 180
OY 181 TGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 181 TGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
OY 241 ATGACCCCTGTGGTGGCTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 241 ATGACCCCTGTGGTGGCTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
OY 301 GGTGCTGGGAGGCTGCGGAGAAATATCTGCTTCCCTCAACTTTTCTGTGGCACATC 360
DB 301 GGTGCTGGGAGGCTGCGGAGAAATATCTGCTTCCCTCAACTTTTCTGTGGCACATC 360
OY 361 GTGCTCATCTTCTCTGAGAGTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GTGCTCATCTTCTCTGAGAGTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
OY 421 GTGAGGAGCGGATTCGCGGAGTCTTCGAGAGCAATCAAGTCTACCGGAGCATATC 480
DB 421 GTGAGGAGCGGATTCGCGGAGTCTTCGAGAGCAATCAAGTCTACCGGAGCATATC 480
```

```

QY 481 GATCTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGCG 540
DB 481 GATCTGCAAAACCTCATGACTCCCTTCAGAAAGCTAACAGTGTGTGGCGCATATGCG 540
QY 541 CCTGAAGACTGGACCTTAACGCTACTTCAATTGACAGCGGTGCCAGTACAGCCGAGAG 600
DB 541 CCTGAAGACTGGACCTTAACGCTACTTCAATTGACAGCGGTGCCAGTACAGCCGAGAG 600
QY 601 AAGTACGGGGGTCCTCTCTGCTGGGCTGCACATCTCTGGCAAAAAGTTGTAAACACA 660
DB 601 AAGTACGGGGGTCCTCTCTGCTGGGCTGCACATCTCTGGCAAAAAGTTGTAAACACA 660
QY 661 CAGTGTGATATGATATGATGAGTATCAGTCAAGTGAAGACAAGTGGATGATCCATCTTACG 720
DB 661 CAGTGTGATATGATATGATGAGTATCAGTCAAGTGAAGACAAGTGGATGATCCATCTTACG 720
QY 721 AAAGGCTCATATCCAGGCGCTGAAAAGCTGCTCCCGGGAACATTTACATTTGTGGTGGC 780
DB 721 AAAGGCTCATATCCAGGCGCTGAAAAGCTGCTCCCGGGAACATTTACATTTGTGGTGGC 780
QY 781 GTCTTCATGCGCATCTGCTGCTGTCAGATATTTGGCANTCTTCCGCAAGACGCTGATC 840
DB 781 GTCTTCATGCGCATCTGCTGCTGTCAGATATTTGGCANTCTTCCGCAAGACGCTGATC 840
QY 841 TCAGACATCGAGGAGTGAAGCGCGCCATCATTTCTGAGAGAGAGATTGAGGAGAGCG 900
DB 841 TCAGACATCGAGGAGTGAAGCGCGCCATCATTTCTGAGAGAGAGATTGAGGAGAGCG 900
QY 901 AGCTGAGCCAGCTGGGAGGCGCAAGCCCTTCTCTGCCATCAGCCCTACGTCCAGAGGGA 960
DB 901 AGCTGAGCCAGCTGGGAGGCGCAAGCCCTTCTCTGCCATCAGCCCTACGTCCAGAGGGA 960
QY 961 GAGGAGCGGACACCCCGGAGGCGAGTCCCATCTTAAGCATACGCGT -GACGTGACCTC 1019
DB 961 GAGGAGCGGACACCCCGGAGGCGAGTCCCATCTTAAGCATACGCGT -GACGTGACCTC 1020
QY 1020 TCTGTTTC -TGCTTGTGCTGCTGCTGAGACCAAGGGTCCCTTGTACTCTCCCAAACTT 1078
DB 1020 TCTGTTTC -TGCTTGTGCTGCTGCTGAGACCAAGGGTCCCTTGTACTCTCCCAAACTT 1080
QY 1079 GTGACTGATCCCTCTGTGATCTTACCCAGAGACAGAGATGTGCTTTATGTGGAGTGG 1138
DB 1081 GTGACTGATCCCTCTGTGATCTTACCCAGAGACAGAGATGTGCTTTATGTGGAGTGG 1140
QY 1139 TGACTCTGAAGACAGAGAGGGCTCTGTGG-----CTGCCAGAGAGGGCTTGACTCAGAC 1193
DB 1141 TGACTCTGAAGACAGAGAGGGCTCTGTGG-----CTGCCAGAGAGGGCTTGACTCAGAC 1200
QY 1194 CCCCTGACAGCTCAACGATGCTGCGAGACACCCCTGCTCCCTGTCGATGGTCCAGAG 1253
DB 1201 CCCCTGACAGCTTAAAGCATGTCTGCAAGACACCTGG -GCCCTCTCCACTGCGATCCAGA 1259
QY 1254 CATCTGCTTTGGGTCATCCACATCTGTGGTGGGCTGTAGAGGAGCCACAGGGGT 1313
DB 1260 CATCTGCTTTGGGTCATCCACATCTGTGGTGGGCTGTAGAGGAGCCACAGGGGT 1319
QY 1314 GGACAGGGGCACTCTCTCTCATCAAGCAAGACGATGGGGGGCTTGGCCGTAACGGGAG 1373
DB 1320 GGACAGGGGCACTCTCTCTCATCAAGCAAGACGATGGGGGGCTTGGCCGTAACGGGAG 1379
QY 1374 GCGNGACGTTGGGCC 1388
DB 1380 GCGNGACGTTGGGCC 1394

```

RESULT 6
 AAZ65350
 ID AAZ65350 standard; DNA; 1667 BP.
 XX
 AC AAZ65350;
 XX
 DT 23-MAR-2000 (first entry)
 XX

DE Human secreted protein gene 10 fragment.
 XX
 KW Human: secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy; chromosome 10; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0958660-A1.
 XX
 PD 18-NOV-1999.
 XX
 PF 06-MAY-1999; 99MO-US09847.
 XX
 PR 12-MAY-1998; 98US-0085093.
 PR 12-MAY-1998; 98US-0085094.
 PR 12-MAY-1998; 98US-0085105.
 PR 12-MAY-1998; 98US-0085180.
 PR 18-MAY-1998; 98US-0085906.
 PR 18-MAY-1998; 98US-0085920.
 PR 18-MAY-1998; 98US-0085921.
 PR 18-MAY-1998; 98US-0085922.
 PR 18-MAY-1998; 98US-0085923.
 PR 18-MAY-1998; 98US-0085924.
 PR 18-MAY-1998; 98US-0085928.
 PR 18-MAY-1998; 98US-0085925.
 PR 18-MAY-1998; 98US-0085927.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 PI Lafleur DW, Endress GA, Ebner R;
 XX
 DR WPI: 2000-062296/05.
 XX
 PT P-PSDB; AAY76266.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -
 XX
 PS Disclosure: Page 427-428; 475pp: English.
 XX
 CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
 CC This sequence was found to be present on human chromosome 10.
 CC AAY76124 to AAY76223 represent the secreted proteins encoded by the 97
 CC human genes. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 97 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
 CC secreted proteins.
 CC
 SQ Sequence 1667 BP; 323 A; 469 C; 488 G; 387 T; 0 other;
 XX

Query Match 85.7%; Score 1189.6; DB 21; Length 1667;
 Best Local Similarity 95.8%; Pred. No. 3e-303;
 Matches 1309; Conservative 1; Mismatches 41; Indels 16; Gaps 8;

|||||
Db 301 ATCTTCTCCGTGAGCTGGCTGTGGCCGTGCTTCTCTCTCCAGACAGCTGGGTAGG 360
Qy 427 GACCGGTTCGGGGAGTCTCTCGAGACATCAACATCTCTACCGGAGATATGATCTG 486
Db 361 GACCGGTTCGGGGAGTCTCTCGAGACATCAACATCTCTACCGGAGATATGATCTG 420
Qy 487 CAAACCTCATGCACTCCCTTCAGAAAGCTAACCAAGTGTGTGGCATATGGCCCTGAA 546
Db 421 CAAACCTCATGCACTCCCTTCAGAAAGCTAACCAAGTGTGTGGCATATGGCCCTGAA 480
Qy 547 GACGTGGACCTCAACGCTCTACTTCATTCAGAGCGGTGCCAGCTAACAGCCGAGAAAGTGC 606
Db 481 GACGTGGACCTCAACGCTCTACTTCATTCAGAGCGGTGCCAGCTAACAGCCGAGAAAGTGC 540
Qy 607 GGGGTCCCTTCTCTGCTGCTGCTGCGAGATCCCTGGCAAAAGTTGTGAACACAGTGT 666
Db 541 GGGGTCCCTTCTCTGCTGCTGCTGCGAGATCCCTGGCAAAAGTTGTGAACACAGTGT 600
Qy 667 GGATATGATGTCAAGGATTCAGCTGAAGAGCAAGTGGGATGATCCATCTTCACGAAAGGC 726
Db 601 GGATATGATGTCAAGGATTCAGCTGAAGAGCAAGTGGGATGATCCATCTTCACGAAAGGC 660
Qy 727 TGCATCCAGCGCTGGAAGAGCTGGCTCCCGGAAACATTTACATTTGGTGGCTCTTC 786
Db 661 TGCATCCAGCGCGGTGGAAGAGCTGGCTCCCGGAAACATTTACATTTGGTGGCTCTTC 720
Qy 787 ATCCGCATCTGCTGCTGTGCGAGATATTTGGCATCTTCTCTGGCAAGACCTGATCTCAGAC 846
Db 721 ATCCGCATCTGCTGCTGTGCGAGATATTTGGCATCTTCTCTGGCAAGACCTGATCTCAGAC 780
Qy 847 ATCGAGCAGTGAAGCGCGGCATCCTCTGA 879
Db 781 ATCGAGCAGTGAAGCGCGGCATCCTCTGA 813

RESULT 11
AAE90658 standard; cDNA; 813 BP.
XX AAE90658;
AC 04-MAY-2001 (first entry)
DT 04-MAY-2001 (first entry)
XX
DE Human TANGO 339 A40V variant cDNA ORF, SEQ ID NO:140.
XX
KW Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW activity modulator; expression modulator; cancer; immunological disorder;
KW cytosolic; immunomodulatory; gene therapy; ss.
XX
OS Homo sapiens.
XX
FN W020109162-A2.
XX
PD 08-FEB-2001.
XX
PE 31-JUL-2000; 2000MO-US20935.
XX
PR 30-JUL-1999; 99US-0365164.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PI Fraser CC, Sharp JD, KIRST SJ, Barnes TM, WRIGHT N, Myers PS;
PI Pan Y;
XX
DR WPI: 2001-138647/14.
DR P-PSDB: AAB87136.
XX
PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
PT disorders -
XX
PS Claim 2; Page 282; 332pp; English.

XX
CC The invention relates to novel secreted/transmembrane proteins, and
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC TANGO 353, TANGO 358, TANGO 365, TANGO 369, TANGO 383, TANGO
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC murine TANGO 393 is also included within the scope of the invention. The
CC invention also encompasses fragments and variants of the proteins of the
CC invention, and nucleic acids encoding them. The invention additionally
CC relates to host cells comprising a nucleic acid of the invention; methods
CC for the production of a protein of the invention; an antibody specific
CC for a protein of the invention; methods for detecting a protein or
CC nucleic acid of the invention; and methods of identifying agents which
CC bind to or modulate the activity of a protein of the invention. The novel
CC secreted proteins, nucleic acids encoding them, and antibodies against
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression or activity of the secreted
CC proteins. The secreted proteins of the invention may also be used to
CC identify modulators of expression or activity, which may be used to
CC the treatment of disorders associated with the proteins of the
CC invention e.g., cancers and immunological disorders. The present
CC sequence represents cDNA encoding human TANGO protein of the invention.
CC
XX
SQ Sequence 813 BP; 163 A; 212 C; 235 G; 203 T; 0 other;
Query Match 58.3%; Score 809.8; DB 22; Length 813;
Best Local Similarity 99.8%; Pred. No. 2.7e+203;
Matches 811; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 67 ATGCACTATATAGATACCTTAACGCCAAGAGCTGCTGTACAAATACCTCTTTTC 126
Db 1 ATGCACTATATAGATACCTTAACGCCAAGAGCTGCTGTACAAATACCTCTTTTC 60
Qy 127 AGCTACAAATACCTCTGCTGTGAGAGTGTCTCTCTGAGAGTGGGCTGGGCA 186
Db 61 AGCTACAAATACCTCTGCTGTGAGAGTGTCTCTCTGAGAGTGGGCTGGGCA 120
Qy 187 TGGAGCGAAAGAGGTGTGCTGCTCGACCTCACCAGAGTACCAGATGGAATGAC 246
Db 121 TGGAGCGAAAGAGGTGTGCTGCTCGACCTCACCAGAGTACCAGATGGAATGAC 180
Qy 247 CCTGTGAGTGTGCTGCTGAGATGAGTGGGCTGTGATGATGACCTGGGCTGGGCTG 306
Db 181 CCTGTGAGTGTGCTGCTGAGATGAGTGGGCTGTGATGATGACCTGGGCTGGGCTG 240
Qy 307 GTGGGGCTGTGCGGAGAAATATCTGCTGCTCACTTTTCTGTGGACCATCTGCTC 366
Db 241 GTGGGGCTGTGCGGAGAAATATCTGCTGCTCACTTTTCTGTGGACCATCTGCTC 300
Qy 367 ATCTCTCTGAGAGCTGTGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db 301 ATCTCTCTGAGAGCTGTGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 427 GACCGGTTCGGGAGTCTCTCGAGAGCAACATCAAGTCTTACCGGAGCATATGATCTG 486
Db 361 GACCGGTTCGGGAGTCTCTCGAGAGCAACATCAAGTCTTACCGGAGCATATGATCTG 420
Qy 487 CAAACCTCATGCACTCCCTTCAGAAAGCTAACCAAGTGTGTGGCATATGGCCCTGAA 546
Db 421 CAAACCTCATGCACTCCCTTCAGAAAGCTAACCAAGTGTGTGGCATATGGCCCTGAA 480
Qy 547 GACTGGAGCTCAACGCTCTACTTCATTCAGAGCGGTGCCAGCTAACAGCCGAGAAAGTGC 606
Db 481 GACTGGAGCTCAACGCTCTACTTCATTCAGAGCGGTGCCAGCTAACAGCCGAGAAAGTGC 540
Qy 607 GGGGTCCCTTCTCTGCTGCTGCTGCGAGATCCCTGGCAAAAGTTGTGAACACAGTGT 666
Db 541 GGGGTCCCTTCTCTGCTGCTGCTGCGAGATCCCTGGCAAAAGTTGTGAACACAGTGT 600
Qy 667 GGATATGATGTCAAGGATTCAGCTGAAGAGCAAGTGGGATGATCCATCTTCACGAAAGGC 726
Db 601 GGATATGATGTCAAGGATTCAGCTGAAGAGCAAGTGGGATGATCCATCTTCACGAAAGGC 660
Qy 727 TGCATCCAGCGCTGGAAGAGCTGGCTCCCGGAAACATTTACATTTGTGGCTCTTC 786

PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YF;
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG16480.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 1; SEQ ID No 16471; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptides and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX Sequence 601 BP; 107 A; 168 C; 180 G; 146 T; 0 other;

Query Match 42.5%; Score 589.8; DB 23; Length 601;
 Best Local Similarity 98.8%; Pred. No. 2.3e-145;
 Matches 594; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 OY 1 CTTCTCGGCCGAGCCGGCGCGCGCTCCCGCCGCGCGGATTCCTTCT 60
 DB 1 CTTCTCGGCCGAGCCGGCGCGCGCTCCCGCCGCGCGGATTCCTTCT 60
 OY 61 CAGAGATGACACTTATATAGTACTCTAACGCCCAAGTACGTGTGATGACAGTAC 120
 DB 61 CAGAGATGACACTTATATAGTACTCTAACGCCCAAGTACGTGTGATGACAGTAC 120
 OY 121 CTTTCAAGTACAAATCATCTTCTGTTGGCTGAGATTGTCTCTTGGAGTCGGGCTG 180
 DB 121 CTTTCAAGTACAAATCATCTTCTGTTGGCTGAGATTGTCTCTTGGAGTCGGGCTG 180
 OY 121 TGGGATGAGACCAAAAGGATGCTGCTCGACCTCACCAGTACCCGGATGATGGA 240
 DB 121 TGGGATGAGACCAAAAGGATGCTGCTCGACCTCACCAGTACCCGGATGATGGA 240
 OY 241 ATCCAGCCTGTGTGCTGCTGATGAGTGGGCGTGTGATGTACCTTGAGGTTCCGC 300
 DB 241 ATCCAGCCTGTGTGCTGCTGATGAGTGGGCGTGTGATGTACCTTGAGGTTCCGC 300
 OY 301 GGGTCGCGGGGGCTGCTGCGGGAATATCTGCTTCACTTTTCTGCGACCATC 360
 DB 301 GGGTCGCGGGGGCTGCTGCGGGAATATCTGCTTCACTTTTCTGCGACCATC 360
 OY 361 GTGCTCATCTTCTCTGAGCTGCGTGGCTGCTGCTCTCTTCTGTTCCAGAGATG 420
 DB 361 GTGCTCATCTTCTCTGAGCTGCGTGGCTGCTGCTCTCTTCTGTTCCAGAGATG 420
 OY 421 GTGAGGACCGGCTTCGGGAGTCTTCGAGACACATCAAGTCTCAACCGGAGCATATC 480

DB 421 GTGAGGACCGGCTTCGGGAGTCTTCGAGACACATCAAGTCTCAACCGGAGCATATC 480
 OY 481 GATCTGCAGAAAACCTTCATGACTCCCTTCAGAAAAGCTAACCATGCTGTGGCCATATG 540
 DB 481 GATCTGCAGAAAACCTTCATGACTCCCTTCAGAAAAGCTAACCATGCTGTGGCCATATG 540
 OY 541 CCTGAGACCTGGGACCTCAACCTTACTTCAATTCAGACGCGGTCCAGCTACAGCCGAGAG 600
 DB 541 CCTGAGACCTGGGACCTCAACCTTACTTCAATTCAGACGCGGTCCAGCTACAGCCGAGAG 600
 OY 601 A 601
 DB 601 A 601
 DB 601 A 601
 RESULT 15
 AAA60182
 ID AAA60182 standard; DNA; 639 BP.
 XX
 XX AAA60182;
 AC
 XX
 XX 02-FEB-2001 (first entry)
 DT
 XX
 DE Hydrophobic domain protein cDNA HP02956 isolated from KB cells.
 XX
 KW Human; secreted protein; membrane protein; hydrophobic domain;
 KW proliferation control; differentiation induction; material transport;
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200029448-A2.
 PD
 XX 25-MAY-2000.
 XX
 PF 17-NOV-1999; 99WO-JP06412.
 XX
 PR 17-NOV-1998; 98JP-0326255.
 XX
 PR 22-DEC-1998; 98JP-0364315.
 PR 16-MAR-1999; 99JP-0069811.
 PR 27-APR-1999; 99JP-0119299.
 PR 19-MAY-1999; 99JP-0138169.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2000-387753/33.
 DR P-PSDB; AAB12120.
 PT
 XX
 XX proteins comprising hydrophobic regions, such as secretory and membrane
 PT proteins, useful in research and diagnostics and having various
 PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
 PT hemostatic, thrombolytic -
 PS Claim 3; Page 195; 410pp; English.
 XX
 CC Secretory proteins play important roles in the proliferation control, the
 CC differentiation induction, the material transport and the biophylaxis of
 CC cells. Membrane proteins have important roles as signal receptors, ion
 CC channels and transporters. The present sequence is the coding sequence
 CC for a human protein which has at least one hydrophobic domain. The
 CC protein encoded by the present sequence may be a secretory or a membrane
 CC protein. The encoded protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or suppressing
 CC activity, haematopoiesis activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 19, 2002, 05:19:27 : Search time 259 Seconds
(without alignments)
2347.643 Million cell updates/sec

Title: US-09-905-674-2
Perfect score: 1449
Sequence: 1 MHYRYSNAKYSWCWKYLLF.....IFLARTLISDIEAVKAGHHF 270

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO/US09905674/runat_16122002.111845_6745/app_query.fasta.1.455
-DB=N.Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsom2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905674 @cgn.1.1.125 @runat.16122002.111845_6745 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1449	100.0	813	22	AAF90629	Human TANGO 339 CD
2	1449	100.0	1388	24	AA29623	Human tetraspan pr
3	1449	100.0	2715	22	AAF90628	Human TANGO 339 CD
4	1446	99.8	813	22	AAF90656	Human TANGO 339 K1
5	1446	99.8	813	22	AAF90657	Human TANGO 339 F2
6	1445	99.7	813	22	AAF90658	Human TANGO 339 A4
7	1445	99.7	813	22	AAF90659	Human TANGO 339 D6
8	1436	99.1	2672	21	AA65259	Human secreted pro
9	1425	98.3	2623	23	AA592016	DNA encoding novel
10	1419	97.9	1428	21	AA93623	Human tetraspanin-
11	1375	94.9	1667	21	AA65350	Human secreted pro
12	1120.5	77.3	639	21	AA60182	Hydrophobic domain
13	1120.5	77.3	2367	21	AAA60192	Hydrophobic domain
14	1049	72.4	816	21	AA93675	Human tetraspanin-
15	936	64.6	601	23	AA80667	DNA encoding novel
16	867	59.8	1655	21	AA98195	Human colon cancer
17	858	59.2	1126	22	AA90014	Clone HNTMH27 codi
18	858	59.2	2091	21	AA96482	CDNA encoding a hu
19	858	59.2	3185	24	AA172633	Human 23228 CDNA.
20	850	58.7	1644	24	AB190838	Human polynucleoti
21	841.5	58.1	2538	24	AA21883	Human TM4SF recept
22	829	57.2	1174	22	AA90015	Human HEBEL16 codi
23	820.5	56.6	1178	22	AA90020	Clone HEBEL16 codi
24	729.5	50.3	1110	22	AA20854	Polynucleotide seq
25	729.5	50.3	1110	22	AA59291	Human CDNA encodin
26	729.5	50.3	1110	24	ABA90960	Human polynucleoti
27	720	49.7	1988	24	ABK35735	CDNA sequence #126
28	681.5	47.0	1214	23	AA58068	DNA encoding novel
29	568	39.2	852	22	AA54553	Human CDNA encodin
30	568	39.2	1932	21	AA98047	Human secreted pro
31	568	39.2	1932	22	AA21660	Human secreted pro
32	568	39.2	1932	24	AA69756	Human secreted pro
33	567	39.1	864	24	AA172287	NEF-4 antisense mo
34	557	38.4	810	23	AB114835	Drosophila melanog
35	550	38.0	2638	24	AB190058	Drosophila melanog
36	541	37.3	1187	23	AB107929	Human polynucleoti
37	522.5	36.1	3244	23	AB107928	Drosophila melanog
38	510	35.2	1130	23	AB116798	Drosophila melanog
39	508	35.1	3009	23	AB114834	Drosophila melanog
40	485	33.5	839	22	AAH34927	Human colon cancer
41	433	29.9	3303	23	AB116798	Drosophila melanog
42	412.5	28.5	1725	22	AA90017	Clone HOFAD65 codi
43	404.5	27.9	1694	19	AAV33190	Secreted protein B
44	403.5	27.8	1772	20	AA200457	Human secreted pro
45	396	27.3	1749	22	AA90023	Clone HORAD65 codi

ALIGNMENTS

RESULT 1	AAF90629	standard; cDNA; 813 BP.	
XX	AAF90629:		
AC	04-MAY-2001	(first entry)	
XX			
DT			
XX			
DE	Human TANGO 339	cdNA ORF, SEQ ID NO:2.	
KX			
KW	Secreted protein; transmembrane protein; TANGO; human; drug screening;		
KW	activity modulator; expression modulator; cancer; immunological disorder;		
KW	cytostatic; immunomodulatory; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WC200109162-A2.		
XX			

PD 08-FEB-2001.
 XX 31-JUL-2000; 2000MO-US20935.
 PF
 XX 30-JUL-1999; 99US-0365164.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;
 PI Pan Y;
 XX WPI; 2001-138647/14.
 DR P-PSDB; AAB87034, AAB87035, AAB87036.
 XX
 PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
 for the prevention, diagnosis and treatment of, e.g. cancers and immune
 disorders -
 PS
 XX Claim 2: Page 220; 332pp; English.

CC The invention relates to novel secreted/transmembrane proteins, and
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
 CC murine TANGO 393 is also included within the scope of the invention. The
 CC invention also encompasses fragments and variants of the proteins of the
 CC invention, and nucleic acids encoding them. The invention additionally
 CC relates to host cells comprising a nucleic acid of the invention; methods
 CC for the production of a protein of the invention; an antibody specific
 CC for a protein of the invention; methods for detecting a protein or
 CC nucleic acid of the invention; and methods for identifying agents which
 CC bind to or modulate the activity of a protein of the invention. The novel
 CC secreted proteins, nucleic acids encoding them, and antibodies against
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression or activity of the secreted
 CC proteins. The secreted proteins of the invention may also be used to
 CC identify modulators of expression or activity, which may be useful in
 CC the treatment of disorders associated with the proteins of the
 CC invention e.g., cancers and immunological disorders. The present
 CC sequence represents cDNA encoding human TANGO protein of the invention.
 XX
 XX Sequence 813 BP; 163 A; 213 C; 235 G; 202 T; 0 other;

Alignment Scores:
 Pred. No.: 5,72e-156 Length: 813
 Score: 1449.00 Matches: 270
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-905-674-2 (1-270) x AAF90629 (1-813)

QY 1 MethISTYTYRArgTYrSerAsnAlaLysValSerCysTIRTYrLysTYrLeuLeuPhe 20
 DB 1 ATCCACTATTATGATGACTCTTAAGCCAGGCTGCTGTGATACAGTACCTCTTTTC 60
 QY 21 SerTYAsnIleIlePheThrLeuAlaGlyValAlaPheLeuGlyValAlaGlyLeuTrpAla 40
 DB 61 AGCTTACACATCAATCTTCTGCTGCTGAGATTGTTCTTCTTGAGATCGGGCTGTGGGCA 120
 QY 41 TTPSerGlnLysGlyValLeuSerAspLeuThrLysValThrArgMethIleSgLYlleasp 60
 DB 121 TGGAGCGAAAGGGGTGCTGTCCGACCTCACCAAGTAGTACCGGATGATGAAATCGAC 180
 QY 61 ProValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
 DB 181 CCTGTGTGTGTCTCTGATGTGGGCTGTGGATGTTCACTCGGGGTTCCTCCCGCTGTC 240
 QY 81 ValGlyAlaLeuArgGlnAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 DB 241 GTGGGGGCTGTGGGAGAAATATCTGCTCAACTTTTCTGTGGCACCATCGTGTCTC 300

QY 101 IlePhePheLeuGlnLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120
 DB 301 ATCTTCTTCCGTGAGAGCTGGGTGGTGGCCGTGGCTCTTCTGTTCCAGGACTGGGTGAG 360
 QY 121 AspArgPheArgGluPhePheGlnSerAsnIleLysSerTYrArgAspIleAspLeu 140
 DB 361 GACCGGTTCCGGGAGTTCTTCGAGACACATCAAGTCTACCGGGACGATATCGATCTG 420
 QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGsgLYAlaTYrGlyProGlu 160
 DB 421 CAAAACTCATGACTCTCTTCAGAAAAGCTTACCAAGTCTGTGGCGATATGGCCCTCGAA 480
 QY 161 AspTrpAspLeuAsnValTYrPheAsnCysSerGlyAlaSerTYrSerArgGluLysCys 180
 DB 481 GACGTGGAGCCTCAACGCTACTTCAATTGACGGGTGCCAGCTACACCGAGAAAGTGC 540
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200
 DB 541 GGGGTCCCTCTCTCTGCTGTGCCAGATCTGCGCAAAAGTTGTGAACACACACTGT 600
 QY 201 GlyTYrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220
 DB 601 GGATATGATGTGCAGGATTCAGCTTAAGAGCAAGTGGGATGATCTCATCTTCACGAAAGGC 660
 QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTYrIleValAlaGlyValPhe 240
 DB 661 TGCATTCAGAGGCGCTGGAAGGTGGCTCCCGGGAACTTATACATGTGGCTGGCGTCTTC 720
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
 DB 721 ATCCCATCTCTGCTGTGCAGATATTGTGGCATCTCTCTGCGCAAGGACGTGATCTAGAC 780
 QY 261 IleGlnAlaValLysAlaGlyHisHisPhe 270
 DB 781 ATCGAGCGAGTGAAGCGCGGCATCACTTC 810

RESULT 2
 AAD29623
 ID AAD29623 standard; cDNA; 1388 BP.
 AC AAD29623;
 XX 17-MAY-2002 (first entry)
 DE Human tetraspan protein, TSPAN-7 encoding cDNA.
 KW Human; tetraspan protein; TSPAN-7; hyperproliferative disorder;
 KW neoplastic disease; prostate cancer; antisense therapy; tumour; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 67..879
 XX /*tag= a
 XX /product= "Human TSPAN-7 protein"
 XX misc_feature 1234..1237
 XX /*tag= b
 XX /note= "These bases are given as YXX in the sequence
 shown as SEQ ID NO:1 in Fig 2 of the specification"
 XX
 XX WO200206340-A2.
 XX 24-JAN-2002.
 XX
 XX 13-JUL-2001; 2001MO-US22134.
 XX
 XX 14-JUL-2000; 2000US-218280P.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Reinhard CJ, Garcia PD;
 XX WPI; 2002-179783/23.

DR P-PSDB: AAE18535.
 XX New inhibitor of tetraspan protein useful for reducing the expression
 PT of activities of tetraspan protein in cells, and for treating a
 PT hyperproliferative disorder, especially cancer in a mammal
 XX
 XX Claim 2; Page 84-85; 88pp; English.
 XX
 CC The invention relates to inhibitors of tetraspan protein, TSPAN-7
 CC which are used to reduce or decrease the expression of TSPAN-7
 CC in a mammalian cell and for treating hyperproliferative disorder,
 CC especially cancer in a mammal. The invention also provides TSPAN-7
 CC polypeptides and polynucleotides. TSPAN-7 inhibitors are useful for
 CC inhibiting proliferation of mammalian cells, including tumour cells,
 CC for decreasing the side effects of cancer therapy and for treating
 CC neoplastic diseases. They are used to modulate TSPAN-7 expression
 CC and function in cancer cells, particularly in prostate cancer cells.
 CC They are also used in antisense therapy. The present sequence is
 CC human TSPAN-7 protein encoding cDNA.
 XX
 SQ Sequence 1388 BP; 270 A; 390 C; 411 G; 313 T; 4 other;
 XX
 Alignment Scores:
 Pred. No.: 1,24e-155 Length: 1388
 Score: 1449.00 Matches: 270
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-905-674-2 (1-270) x AAD29623 (1-1388)
 QY 1 METHISTYRTYRARGTYISERASNAIALYSVALISERCYSTRTPTRYLTYRILEUENPHE 20
 DB 67 ATGCACATATATAGTACTCTACCCCAAGTCACCTGCTGATCAAGTACCTCTTTTC 126
 QY 21 SERTYRASNILEILEPHEPTRLEUALAGLYVALPHELEGLYVALGLYLEUTRPA 40
 DB 127 ACCTACACATCATCTCTGCTGGTGGCTGAGTGTCTTCCCTTGAGTGGGGCTGGGCA 186
 QY 41 TRPSERGLULYSGLYVALLEUSERASPLEUTRILYSVALINRARGMETHISGLYLEASP 60
 DB 187 TGGAGGGAAGGGTGTGTCTGCACCTCACCAAGTACCCGGATCATGATTCAGAC 246
 QY 61 PROVALLEULEUVALLEUMETVALGLYVALIMETPHEPTRLEUGLYPHEALAGLYYS 80
 DB 247 CCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
 QY 81 VALGLYALALEUALRGYUASNILLECYSLLEULASNPHEPCYGLYTHRIIEVALLEU 100
 DB 307 GTGGGGGCTCTCGGGAGAAATATGCTTGCTCAACTTTTCTGTGGCACCATCTGCTTC 366
 QY 101 ILEPHEPHELEUGLULEUALVALALAVALLEUALPHELEUPHEGLINASPTRYVALRG 120
 DB 367 AACTCTTCTTCTGAGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGG 426
 QY 121 ASPATGPHEARGLUPHEPHEGLUSERASNILEYSSERTYRARGASPSPILEASPLEU 140
 DB 427 GACCGGTCGGGAGATTCTTCAGAGACACATCAAGTCTTACCGGGAAGATATCATCTG 486
 QY 141 GLINASLEULEASPSERLEUGLINSALASNGLNCYSGLYALATYRGLYPROGLU 160
 DB 487 CAACACCTCATGACTCCCTTCAGAAAGCTAACAGTGTGGGGCATATGGCCCTAA 546
 QY 161 ASPTTPASPLEUASNVALLTYRHEASNCYSSERGLYALASERTYRSEARGLULYSCYS 180
 DB 547 GACTGGGACTTCAACGCTACTTCAATTCACGCGGTGCAGTACAGCGGAGAGAAATGTC 606
 QY 181 GLYVALPROPHESERCYSCYVALPROASPPROAGLINSVALVALASPTHGLINYS 200
 DB 607 GGGGTCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
 QY 201 GLTYRASPVALARGILEGLINLEULYSSERYSTRPASGLUSERILEPHEPTRLYSGLY 220

DB 667 GGATATGATGTCCAGATTACACTGACAGCAAGGGGATGATGCTCATCTTCAGAAAGGC 726
 QY 221 CYSILLEGIALALEUGLUSERTRIPLEUPROARGASNILEYTRILEVALALAGLYVALPHE 240
 DB 727 TGCATCCAGGCGCTGGAAGAGTGTCTCCCGGAACTTATACATTTGCTGGCTGCTTTC 786
 QY 241 ILEALATSESRLEULEUGLILEPHEGLYILEPHELEUALARGPTRLEULISERASP 260
 DB 787 ATCGCATCTGCTGCTTGTGCAGATATTTGGCATCTTCTTGGCAAGAGCTGATCTCAGAC 846
 QY 261 ILEGLUALVALYLSALAGLYHSHISAPHE 270
 DB 847 ATCAGAGCAGTGAAGGCGCGGCATCACTTC 876
 RESULT 3
 AAF90628
 ID AAF90628 standard; cDNA; 2715 BP.
 XX
 AC AAF90628;
 XX
 DT 04-MAY-2001 (first entry)
 XX
 DE Human TANGO 339 cDNA, SEQ ID NO:1.
 XX
 KW Secreted protein; transmembrane protein; TANGO; human; drug screening;
 KW activity modulator; expression modulator; cancer; immunological disorder;
 KW cyostatic; immunomodulatory; gene therapy; ss.
 OS Homo sapiens.
 XX
 PN WO200109162-A2
 XX
 PD 08-FEB-2001.
 XX
 PE 31-JUL-2000; 2000WO-US20935.
 XX
 PR 30-JUL-1999; 99US-0365164.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;
 PI Pan Y;
 XX
 DR WPI: 2001-138647/14.
 DR P-PSDB: AAB87034, AAB87035, AAB87036.
 XX
 PS Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
 PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
 PT disorders -
 XX
 PS Claim 2; Page 219-220; 332pp; English.
 XX
 CC The invention relates to novel secreted/transmembrane proteins, and
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
 CC murine TANGO 393 is also included within the scope of the invention. The
 CC invention also encompasses fragments and variants of the proteins of the
 CC invention, and nucleic acids encoding them. The invention additionally
 CC relates to host cells comprising a nucleic acid of the invention; methods
 CC for the production of a protein of the invention; an antibody specific
 CC for a protein of the invention; methods for detecting a protein or
 CC nucleic acid of the invention; and methods of identifying agents which
 CC bind to or modulate the activity of a protein of the invention. The novel
 CC secreted proteins, nucleic acids encoding them, and antibodies against
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression or activity of the secreted
 CC proteins. The secreted proteins of the invention may also be used to
 CC identify modulators of expression or activity, which may be useful in
 CC the treatment of disorders associated with the proteins of the
 CC invention e.g., cancers and immunological disorders. The present
 CC sequence represents cDNA encoding human TANGO protein of the invention.

```

XX      Sequence 2715 BP: 559 A; 760 C; 715 G; 681 T; 0 other:
SQ
Alignment Scores:
Pred. No.:      3,25e-155      Length:      2715
Score:          1449.00      Matches:      270
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             22      Gaps:      0

US-09-905-674-2 (1-270) x AAF90628 (1-2715)
QY      1 MethisTyrTyrArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuPhe 20
DB      210 ATGCACATATTATGATGACTTAACGCCAAGGTACAGTCTGCTGATACAGTACTCTTTTC 269
QY      21 SerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
DB      270 AGCTACACATCATCTTCTGTGGTGGATGATGCTTCTTCTTGAGTGGGCTGTGGCA 329
QY      41 TrpSerGlnLysGlyValLeuSerAspLeuTrpLysValThrArgMethisGlyLeuAsp 60
DB      330 TGGAGCGAAAGGGTGTGTGTCGCCACCTCACCAAGTACCCGGATGATGCAATCGAC 389
QY      61 ProValIleuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCys 80
DB      390 CCGTGTGCTGTGCTCTGATGAGGCGCTGTGATGCTTACCCCTGGGCTGCGCGCTGC 449
QY      81 ValGlyAlaLeuArgGlnAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
DB      450 GTGGGGGCTGTGGGAGATATCTGCTCAACTTTTCTGTGGCCACATCGTCTC 509
QY      101 IlePhePheLeuGlnLeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArg 120
DB      510 ATCTTTCTTCTGGAGCTGTGCTGTGGCGCTGTGGCTTCTCTTCCAGAGACTGGGTGAGG 569
QY      121 AspArgPheArgGlnPhePheGlnSerAsnIleLysSerTyrArgAspIleAspLeu 140
DB      570 GACCGTTCGGGAGTCTTCCAGAGCAACATCAGTCTACCGGGAGCATTCATCTCG 629
QY      141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160
DB      630 CAAACCTCATGCACTCCCTCAGAAAGCTAACCAAGTCTGTGGCGCATATGCCCTGAA 689
QY      161 AspTrpAspLeuAsnValLyrPheAsnCysSerGlyAlaSerTyrSerArgGluLysCys 180
DB      690 GATGGGACCTCAACGCTTACTTCAATTCACGGGTGCGACGCTACAGCCGAGAGAAGTGC 749
QY      181 GlyValProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCys 200
DB      750 GGGGTCCTCTTCCGTCGTCGCGCAAGATCCTGGCGCAAAAGTTGTGAAACACAGTGT 809
QY      201 GlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGlnSerIlePheThrLysGly 220
DB      810 GGATATGATGTCAGGATTCAGCTGAAGACAGTGGATGAGTCACTTCCGAAAGGC 869
QY      221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
DB      870 TGCATTCAGGCGCTGGAAGGTGCTCCCGCGGACATTTACATTTGTGCTGCTTTC 929
QY      241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
DB      930 ATCGCATCTCGCTGTGTCAGATATTGGCATCTTCTCGCAAGGACGCTGATCTCAGAC 989
QY      261 IleGlnAlaValLysAlaGlyHisHisPhe 270
DB      990 ATCGAGGCAGTGAAGCCGCGCATCTTC 1019
RESULT 4
AAF90656
ID      AAF90656 standard; cDNA; 813 BP.
XX

```

```

AC      AAF90656;
XX
XX      04-MAY-2001 (first entry)
DE      Human TANGO 339 K10R variant cDNA ORF, SEQ ID NO:136.
XX
XX      Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW      activity modulator; expression modulator; cancer; immunological disorder;
KW      cytostatic; immunomodulatory; gene therapy; ss.
XX
XX      Homo sapiens.
OS
XX      WO200109162-A2.
XX
XX      08-FEB-2001.
XX
XX      31-JUL-2000; 2000WO-US20935.
XX
XX      30-JUL-1999; 99US-0365164.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
PA
XX      Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wighton N, Myers PS;
PI      Pan Y;
XX
XX      WPI: 2001-138647/14.
DR      P-PSDB; AAB87134.
XX
XX      Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT      for the prevention, diagnosis and treatment of, e.g. cancers and immune
PI      disorders -
XX
XX      Claim 2; Page 279; 332pp; English.
XX
XX      The invention relates to novel secreted/transmembrane proteins, and
CC      nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC      TANGO 353, TANGO 358, TANGO 365, TANGO 369, TANGO 383, TANGO
CC      393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
CC      murine TANGO 393 is also included within the scope of the invention. The
CC      invention also encompasses fragments and variants of the proteins of the
CC      invention, and nucleic acids encoding them. The invention additionally
CC      relates to host cells comprising a nucleic acid of the invention; methods
CC      for the production of a protein of the invention; an antibody specific
CC      for a protein of the invention; methods for detecting a protein or
CC      nucleic acid of the invention; and methods for identifying agents which
CC      bind to or modulate the activity of a protein of the invention. The novel
CC      secreted proteins, nucleic acids encoding them, and antibodies against
CC      them may be used in the prevention, diagnosis and treatment of diseases
CC      associated with inappropriate expression or activity of the secreted
CC      proteins. The secreted proteins of the invention may also be used to
CC      identify modulators of expression or activity, which may be useful in
CC      the treatment of disorders associated with the proteins of the
CC      invention e.g., cancers and immunological disorders. The present
CC      sequence represents cDNA encoding human TANGO protein of the invention.
XX
XX      Sequence 813 BP: 162 A; 213 C; 236 G; 202 T; 0 other:
SQ
Alignment Scores:
Pred. No.:      1,26e-155      Length:      813
Score:          1446.00      Matches:      269
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.63%      Mismatches: 0
Query Match:    99.79%      Indels:      0
DB:             22      Gaps:      0

US-09-905-674-2 (1-270) x AAF90656 (1-813)
QY      1 MethisTyrTyrArgTyrSerAsnAlaLysValSerCysTrpTyrLysTyrLeuPhe 20
DB      1 ATGCACATATTATGATGACTTAACGCCAAGGTACAGTCTGCTGATACAGTACTCTTTTC 60
QY      21 SerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40

```

Db 61 AGCTACACATCATCTTGGTGGCTGGAGTTGCTTCCTTGAGACTCGGAGCTGGGCA 120
 QY 41 TTPSerGIuLysGIValIeuSerAspLeuThrLysValThrArgMetHisGIleAsp 60
 Db 121 TGGAGGAAAGGGCTGTGCTGTCCGACCTCACCAGAAAGTACCAGGATGATGAAATCGAC 180
 QY 61 ProValIValIeuValIeuMetValGIValValMetPheThrLeuGIYpheaIagIcys 80
 Db 181 CCTGTGTGTCTGTGCTGATGGGCGGTGATGTTCACCTGGGGTTCGCCGGCTGC 240
 QY 81 ValGIValIeuValIeuValIeuValIeuValIeuValIeuValIeuValIeuValIeu 100
 Db 241 GTGGGGGCTCTGGGAGATATCTGCTCAACTTTTCTGTGGCCATCTGCTC 300
 QY 101 ILePhePheLeuGIuLeuAlaValAlaValIeuAlaPheLeuPheGIuAspIrrValArg 120
 Db 301 ATCTTCTCTGTGAGCTGTGCTGTGGCCGTGCTGGCTTCTTCTTCAGAGACTGGGTGAGG 360
 QY 121 AspArgPheArgIuPhePheGIuSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
 Db 361 GACCCGCTTCGGGAGATTCTTCGAGACACATCAAGTCTACCGGGAGCATATCATCTG 420
 QY 141 GIuAsnLeuIleAspSerLeuGIuLysAlaAsnGIuLysCysGIValArgIleProGIu 160
 Db 421 CAHAACCTATGACTCCCTTCAAGAAAGCTAACCAAGTGTGGGCAATATGGCCCTGAA 480
 QY 161 AspIrrPheAspLeuAsnValTyrPheAsnCysSerGIYAlaSerTyrSerArgIuLysCys 180
 Db 481 GACTGGAGCTCAAGCTCTACTTCAATTTCACGCGGTGACAGCCGAGAGAAAGTGC 540
 QY 181 GIYValIrrPheSerCysCysValProAspProAlaGIuLysValValAsnThrGIuLys 200
 Db 541 GGGGTCCCTCTCTCTGCTGCGTCAGATCTCTGCGCAAAAGTGTGAAACACACAGTGT 600
 QY 201 GIYTYrAspValArgIleGIuLysSerLysTrrPaspGIuSerIlePheThrIysGI 220
 Db 601 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 221 CysIleGIuAlaLeuGIuSerTrrPleuProArgAsnIleTyrIleValAlaIagIValPhe 240
 Db 661 TGCATCCAGGCGCTGGAAGCTGGCTCCGGGAGAACATTTCATGTGTGCTGGCTTTC 720
 QY 241 ILeAlaIleSerLeuLeuGIuIlePheGIYIlePheLeuAlaArgThrLeuIleSerAsp 260
 Db 721 ATCGGCATCTCTGCTGTGAGATATTTGGCATCTCTCTGCGAAGACGCTGATCTCAGAC 780
 QY 261 ILeGIuAlaValIleValIagIYHisIshPhe 270
 Db 781 ATCGAGGCAGTGAAGCGCGGCATCATCTTC 810
 RESULT 5
 AAF90657
 ID AAF90657 standard; cDNA; 813 BP.
 AC AAF90657;
 XX
 DT 04-MAY-2001 (first entry)
 XX
 DE Human TANGO 339 F20Y variant cDNA ORF, SEQ ID NO:138.
 XX
 KW Secreted protein; transmembrane protein; TANGO; human; drug screening;
 KW activity modulator; expression modulator; cancer; immunological disorder;
 KW cytosolic; immunomodulatory; gene therapy; ss.
 OS Homo sapiens.
 XX
 PN WO200109162-A2.
 XX
 PD 08-FEB-2001.
 XX
 PF 31-JUL-2000; 2000WO-US20935.
 XX
 PR 30-JUL-1999; 99US-0365164.

XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC, Sharp JD, Kirst SJ, Barnes TM, Wrighton N, Myers PS;
 PI Pan Y;
 XX
 DR WPI; 2001-138647/14.
 XX
 PF P-PSDB; AAB87135.
 XX
 PT Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
 PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
 PT disorders -
 XX
 PS Claim 2; Page 280-281; 332pp; English.
 XX
 CC The invention relates to novel secreted/transmembrane proteins, and
 CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
 CC TANGO 353, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
 CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a
 CC murine TANGO 393 is also included within the scope of the invention. The
 CC invention also encompasses fragments and variants of the proteins of the
 CC invention, and nucleic acids encoding them. The invention additionally
 CC relates to host cells comprising a nucleic acid of the invention; methods
 CC for the production of a protein of the invention; an antibody specific
 CC for a protein of the invention; methods for detecting a protein or
 CC nucleic acid of the invention; and methods of identifying agents which
 CC bind to or modulate the activity of a protein of the invention. The novel
 CC secreted proteins, nucleic acids encoding them, and antibodies against
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate expression or activity of the secreted
 CC proteins. The secreted proteins of the invention may also be used to
 CC identify modulators of expression or activity, which may be useful in
 CC the treatment of disorders associated with the proteins of the
 CC invention e.g., cancers and immunological disorders. The present
 CC sequence represents cDNA encoding human TANGO protein of the invention.
 XX
 SQ Sequence 813 BP; 164 A; 213 C; 235 G; 201 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,26e-155 Length: 813
 Score: 1446.00 Matches: 269
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.63% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 22 Gaps: 0
 US-09-905-674-2 (1-270) x AAF90657 (1-813)
 QY 1 MethIstYrTYrArgTYrSerAsnAlaLysValSerCysTrrPYrLysTYrLeuPhe 20
 Db 1 ATGCACATATTTATATATCTCAACGCCAAGGTACAGCTGTGTGTCACAAAGTACTTCCTTTAC 60
 QY 21 SerTYrAsnIleIlePheTrrPleuAlaGIYValValPheLeuGIYValGIYLeuTrrPala 40
 Db 61 AGCTACAAACATCATCTCTGTGCTGGTGGAGTGTCTTCTTCTGGAAGCGGGGTGGGCA 120
 QY 41 TrrSerGIuLysGIYValIeuSerAspLeuThrLysValThrArgMetHisGIYIleAsp 60
 Db 121 TGGAGGAAAGGGGTGTGCTGTGCGACCTCAACAAATGACCGGAGTGTGAATTCGAC 180
 QY 61 ProValIValIeuValIeuMetValGIYValValMetPheThrLeuGIYpheaIagIcys 80
 Db 181 CCTGTGTGTCTGTGCTGATGGGCGGTGATGTTCACCTGGGGTTCGCCGGCTGC 240
 QY 81 ValGIYValIeuValIeuValIeuValIeuValIeuValIeuValIeuValIeuValIeu 100
 Db 241 GTGGGGGCTCTGGGAGATATCTGCTCAACTTTTCTGTGGCCATCTGCTC 300
 QY 101 ILePhePheLeuGIuLeuAlaValAlaValIeuAlaPheLeuPheGIuAspIrrValArg 120
 Db 301 ATCTTCTCTGTGAGCTGTGCTGTGGCCGTGCTGGCTTCTTCTTCAGAGACTGGGTGAGG 360
 QY 121 AspArgPheArgIuPhePheGIuSerAsnIleLysSerTyrArgAspAspIleAspLeu 140

DB 361 GACCGGTTCCGGAGTCTTCTGACAGCAACATCAAGTCTTACCGGAGCATATCATGTG 420
QY 141 GlnAsnLeuIleAspSerLeuGlnIlysaIlaAsnGlnCysGlyAlaIleValProGlu 160
DB 421 CAAACCTCATGACCTCCCTCAGAAAGCTAACAGTCTGCGGCGCATATGGCCCTGAA 480
QY 161 AspTrpAspLeuAsnValIlePheAsnGlySerGlyAlaSerTyrSerArgGlnIlyscys 180
DB 481 GACTGGGACCTCAACGCTTACTTCAATTGACAGCGGTGCGACGTACAGCGGAGAAAGTGC 540
QY 181 GlyAlaProPheSerCysCysValProAspProAlaGlnIlysaValIleAsnThrGlnCys 200
DB 541 GGGGTCCCTTCTCTCTGCGGTGCGCAAGTCTGCGCAAAAAGTGTGACACACAGAGT 600
QY 201 GlyTyrAspValArgIleGlnLeuIlySerIlyTrpAspGluSerIlePheThrIlyscly 220
DB 601 GGATATGATGTCAGATTCAGCTGGAAGAGCAAGTGGATGATGATCCATCTTCAGAAAGCG 660
QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
DB 661 TGCATTCAGGCGCTGAAAGCTGGCTCCCGGGAACATTTCATTTGAGCTGGCTCTTC 720
QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
DB 721 ATCGGCATCTCGCTGTGCGAGATTTGGCATCTTCTGCGCAAGAGACGCTGATCTCAGAC 780
QY 261 IleGluAlaValIlysaIleGlyHisPhe 270
DB 781 ATCGAGGAGTGAAGCGCGCATCATCTTC 810
RESULT 6
ID AAF90658 standard; cDNA; 813 BP.
XX AAF90658;
AC AAF90658;
XX AAF90658;
DT 04-MAY-2001 (first entry)
XX 04-MAY-2001 (first entry)
DE Human TANGO 339 A40V variant cDNA ORF, SEQ ID NO:140.
XX
XX Secreted protein; transmembrane protein; TANGO; human; drug screening;
KW activity modulator; expression modulator; cancer; immunological disorder;
XX cytosolic; immunomodulatory; gene therapy; ss.
OS Homo sapiens.
XX
XX WO200109162-A2.
XX
XX 08-FEB-2001.
XX
XX 31-JUL-2000; 2000WO-US20935.
XX
XX 30-JUL-1999; 99US-0365164.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI Frazer CC, Sharp JD, Kirst ST, Barnes TM, Wrighton N, Myers PS;
PI Pan Y;
XX
XX WPI: 2001-138647/14.
XX
XX P-PSDB: AAB87136.
XX
XX Nucleic acids encoding secreted TANGO and MANGO polypeptides, useful
PT for the prevention, diagnosis and treatment of, e.g. cancers and immune
PT disorders -
XX
XX
XX Claim 2; Page 282; 332pp; English.
XX
XX The invention relates to novel secreted/transmembrane proteins, and
CC nucleic acids encoding them. The novel proteins are designated TANGO 339,
CC TANGO 333, TANGO 358, TANGO 365, TANGO 368, TANGO 369, TANGO 383, TANGO
CC 393, TANGO 402, MANGO 346 and MANGO 349 and are of human origin, and a

CC murine TANGO 393 is also included within the scope of the invention. The
CC invention also encompasses fragments and variants of the proteins of the
CC invention, and nucleic acids encoding them. The invention additionally
CC relates to host cells comprising a nucleic acid of the invention; methods
CC for the production of a protein of the invention; an antibody specific
CC for a protein of the invention; methods for detecting a protein or
CC nucleic acid of the invention; and methods of identifying agents which
CC bind to or modulate the activity of a protein of the invention. The novel
CC secreted proteins, nucleic acids encoding them, and antibodies against
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression or activity of the secreted
CC proteins. The secreted proteins of the invention may also be used to
CC identify modulators of expression or activity, which may be useful in
CC the treatment of disorders associated with the proteins of the
CC invention e.g., cancers and immunological disorders. The present
CC sequence represents cDNA encoding human TANGO protein of the invention.
XX
SQ Sequence 813 BP; 163 A; 212 C; 235 G; 203 T; 0 other;

Alignment Scores:

Pred. No.:	1 64e-155	Length:	813
Score:	1445.00	Matches:	269
Percent Similarity:	99.63%	Conservative:	0
Best Local Similarity:	99.63%	Mismatches:	1
Query Match:	99.72%	Indels:	0
DB:	22	Gaps:	0

US-09-905-674-2 (1-270) x AAF90658 (1-813)

QY 1 MethIstYrYrArgYrSerAsnAlaIlysaIleValSerCysTrpYrIlysclyLeuPhe 20
DB 1 ATGCATATATATAGATACCTTCAACGCCAAGTCAAGTGTGATACAAATCACTCTTTC 60
QY 21 SerTyrAsnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAla 40
DB 61 AGCTACAAACATCATCTCTGCTGGTGGAGTGTCTCTTGGAGTGGGCTGGTGA 120
QY 41 TrpSerGluIlysclyValLeuSerAspLeuThrIlysaIleValThrArgMetHisGlyIleAsp 60
DB 121 TGGAGCGAAAGGCTGTGCTGTGCGACCTCACCAATGACCCGGATGCATGGAAATGAC 180
QY 61 ProValIleLeuValLeuMetValGlyValIlePheThrLeuGlyPheAlaGlyCys 80
DB 181 CCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnIlePheCysGlyThrIleValLeu 100
DB 241 GTGGGGGCTGTGGGGAATATCTGCTCAACTTTTCTGTGGACACATCTGTCTC 300
QY 101 IlePhePheLeuGlnLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120
DB 301 ATCTTCTTCCGTGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
QY 121 AspArgPheArgGluPhePheGlnSerAsnIleIlySerTyrArgAspIleAspLeu 140
DB 361 GACCGGTTCGCGGAGTCTTCTGAGAGCAACATCAAGTCTTACCGGAGCGATTCGATCTG 420
QY 141 GlnAsnLeuIleAspSerLeuGlnIlysaIlaAsnGlnCysGlyAlaIleValProGlu 160
DB 421 CAAACCTCATGACCTCCCTCAGAAAGCTAACAGTGTGTGGCGATATGGCCCTGAA 480
QY 161 AspTrpAspLeuAsnValIlyPheAsnGlySerGlyAlaSerTyrSerArgGlnIlyscys 180
DB 481 GACTGGGACCTCAACGCTTACTTCAATTGACAGCGGTGCGCATACACGAGAAAGTGC 540
QY 181 GlyAlaProPheSerCysCysValProAspProAlaGlnIlysaValIleAsnThrGlnCys 200
DB 541 GGGGTCCCTTCTCTCTGCGGTGCGCAAGTCTGCGCAAAAAGTGTGACACACAGAGT 600
QY 201 GlyTyrAspValArgIleGlnLeuIlySerIlyTrpAspGluSerIlePheThrIlyscly 220
DB 601 GGATATGATGTCAGATTCAGCTGGAAGAGCAAGTGGATGATGATCCATCTTCAGAAAGCG 660

XX Human secreted protein gene 10.
 DE
 XX
 XX Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukemia;
 KW therapy; chromosome 10; ds.
 XX
 OS Homo sapiens.
 XX
 XX MO9958660-A1
 XX 18-NOV-1999.
 XX
 XX 06-MAY-1999; 99WO-US09847.
 XX
 XX 12-MAY-1998; 98US-0085093.
 XX 12-MAY-1998; 98US-0085094.
 XX 12-MAY-1998; 98US-0085105.
 XX 12-MAY-1998; 98US-0085180.
 XX 18-MAY-1998; 98US-0085906.
 XX 18-MAY-1998; 98US-0085920.
 XX 18-MAY-1998; 98US-0085921.
 XX 18-MAY-1998; 98US-0085922.
 XX 18-MAY-1998; 98US-0085923.
 XX 18-MAY-1998; 98US-0085924.
 XX 18-MAY-1998; 98US-0085928.
 XX 18-MAY-1998; 98US-0085925.
 XX 18-MAY-1998; 98US-0085927.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 PI Lafleur DW, Endress GA, Ebner R;
 XX
 XX WPI: 2000-062296/05.
 XX P-PSDB; AAY76133.
 XX
 XX New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -
 XX
 XX Claim 1: Page 301-302; 475pp; English.
 PS
 XX AA265250 to AA265350 represent 97 isolated human secreted protein genes.
 CC This sequence was found to be present on human chromosome 10.
 CC AAY76124 to AAY76223 represent the secreted proteins encoded by the 97
 CC human genes. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 97 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
 CC secreted proteins.
 XX
 SQ Sequence 2672 BP; 548 A; 747 C; 708 G; 661 T; 8 other;

Alignment Scores: 9,73e-154 Length: 2672
 Pred. No.: 1436.00 Matches: 269
 Score:

Percent Similarity: 99.63% Conservative: 0
 Best Local Similarity: 99.63% Mismatches: 1
 Query Match: 99.10% Indels: 0
 DB: 21 Gaps: 0
 US-09-905-674-2 (1-270) x AA265259 (1-2672)
 QY 1 Methistyr-TYArgTyrSerAsnAlaValSerCysTyrPyrLysTyrLeuPhe 20
 DB 178 ATGCATATTATTAAGATACCTTAACGCCAGAGTCAGTGTGTACAGACTCTTTTC 237
 QY 21 SerTyrAsnIleIlePhePheLeuAlaGlyValIlePheLeuGlyValGlyLeuTrrPhe 40
 DB 238 ACCTACACATCATCTTCTGRTTGGCTGAGCTTCTTCCTTGAGTGGCGCTGGGCA 297
 QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrArgMethIstGlyLeuAsp 60
 DB 298 TGGAGCGAAGAGGCTGTGCTGCTCCGACCTCACCAGTACCCGGATGATGATTCGAC 357
 QY 61 ProValValLeuValLeuMetValGlyValIleMetPheThrLeuGlyPheAlaGlyCys 80
 DB 358 CTTGTGTGTGTGTCTCTGATAGTGGCGTGTGTATGTCACCTGGGCTTCGCGCTGC 417
 QY 81 ValGlyAlaLeuArgGlyAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 DB 418 GTGGGGCTCTGGCGAGAAATATCTCTGCTCACTTTTCTGTGGCACCATGCTC 477
 QY 101 IlePhePheLeuGlyLeuValAlaValAlaValIlePheLeuPheGlnAspTrrPheValArg 120
 DB 478 ATCTCTCTCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
 QY 121 AspArgPheArgGlyIlePhePheGlyLeuSerAsnIleLysSerTyrTrpArgAspIleAspLeu 140
 DB 538 GACCGGTTCGGGAGGTCTCTCGAGACCAATCAATCTACCGGACCATATCTGATCTG 597
 QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyrGlyProGlu 160
 DB 598 CAAACCTCATGACCTCTCCAGAAAGCTAACACAGTGGCTGTCATATGGCCCTGAA 657
 QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCys 180
 DB 658 GACTGGAGCTCAACGCTCACTTCAATTCGACGGGTGACGCTACAGCGAGAGAGTGC 717
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnCys 200
 DB 718 GGGGTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
 QY 201 GlyTrpAspValArgIleGlnLeuLysSerLysTrpAspGlyLeuSerIlePheThrLysGly 220
 DB 778 GGATATATGTCAAGATTCAAGTCAAGCAAGTGGGATAGTCCATCTTCACGAAAGGC 837
 QY 221 CysIleGlnAlaLeuGlnSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
 DB 838 TGCATTCAGAGCGCTGGAAGACTGGCTCCCGGAACTTATTAATTTGGCTGCGCTTTC 897
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
 DB 898 ATCCCAATCTTCGCTGTGCAAGATATTTGGCATCTCTCTGCAAGGACGTGATCTAGAC 927
 QY 261 IleGlnAlaValLysAlaGlyLysHisPhe 270
 DB 958 ATCGAAGGAGTGAAGCGCGGCATCACTTC 987
 RESULT 9
 AAS92016
 ID AAS92016 standard; cDNA; 2623 BP.
 AC AAS92016;
 XX 13-FEB-2002 (first entry)
 DT
 XX DNA encoding novel human diagnostic protein #27820.
 DE
 XX

XX 09-MAR-1999; 99US-0123667.
 PR 08-MAR-2000; 2000US-0123667.
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA;
 PI
 XX WPI: 2000-594318/56.
 DR P-PSDB: AAB23036.
 DR
 XX Novel human membrane associated or secreted polypeptides and
 PT polynucleotides useful for diagnosis, prevention and treatment of
 PT pathological states such as cancer, immune, cardiovascular and
 PT neurological disorders
 PS
 XX Claim 3; Fig 8A; 151pp; English.

CC Sequences AAA93616-A93631 and AAA93673-A93676 represent nucleic acids
 CC which encode human SECC proteins (AAB23029-B23048). The SECC proteins
 CC of the invention are either secreted or membrane-associated proteins
 CC and act as regulator of cellular proliferation and differentiation. SECC
 CC proteins or nucleotides are useful for diagnosing the presence of, or
 CC predisposition to, a disease associated with altered levels of SECC
 CC proteins and nucleotides. The SECC proteins are also useful to screen
 CC compounds that modulate SECC activity or expression. The interaction of
 CC a SECC protein with other cellular proteins may be useful to modulate
 CC the activity of a partner protein, cellular proliferation, cellular
 CC differentiation and cell survival. SECC nucleotides are useful for the
 CC recombinant expression of SECC protein, and may be used to detect SECC mRNA
 CC or genetic lesions in the SECC gene. They may also be used to modulate
 CC SECC expression (e.g., using antisense oligonucleotides). SECC nucleic
 CC acid sequences are also useful for identifying a cell or tissue type in
 CC a biological sample, and in forensic biology. SECC primers or probes are
 CC useful for detecting the presence of SECC nucleotides and for screening
 CC tissue cultures for contamination. Diseases that may be treated or
 CC prevented using SECC proteins or nucleotides include cancer (e.g.,
 CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
 CC infections, inflammatory disorders, arthritis, hematopoietic disorders,
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.
 CC
 XX
 XX
 SQ Sequence 1428 BP; 279 A; 408 C; 426 G; 315 T; 0 other;

Alignment Scores:

Pred. No.: 3,47e-152 Length: 1428
 Score: 1419.00 Matches: 265
 Percent Similarity: 98.15% Conservative: 0
 Best Local Similarity: 98.15% Mismatches: 5
 Query Match: 97.93% Indels: 0
 DB: 21 Gaps: 0

US-09-905-674-2 (1-270) x AAA93623 (1-1428)

QY 1 MethistyrtyrargtyrserasnalyvalysercrrpyrlytyrleuPhe 20
 DB 96 ATGCACTATTATGATACCTCTAACGCCAAGGTACACTCTGTACAGTACCTCTTTTC 155
 QY 21 SerTyrAsnIleIlePheThrPheuAaIyValIlePheLeuGlyValGlyLeuTrpAla 40
 DB 156 AGCTACACATCATCTTCTTCGTGGTGGAGTGTCTTCCTTGAGTGGGGTGTGGGCA 215
 QY 41 TrpSerGluGlyGlyValLeuSerAspLeuThrIleValIleThrArgMetHisGlyIleAsp 60
 DB 216 TGGAGCGAAAAGGGTGTGCTCCGCCCTCACCAAGAGGACCCGGATGCATGGAATTCAC 275
 QY 61 ProValValLeuValLeuMetValGlyValIleMetPheThrLeuGlyPheAlaGlyCys 80
 DB 276 CTTGGCGTGTGCTCTCGATGGTGGCGCGGTGATGTTTCACCCCTTGGGGTTCGCCGCGCGC 335

QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
 DB 336 GTGGGGGCGGCGAGGAGATATATGCTGCTCAACTTTTCTGTGGACCATCGTGTCTG 395
 QY 101 IlePhePheLeuGluLeuAaIyValAlaValLeuAlaIlePheLeuPheGlnAspTrpValArg 120
 DB 396 ATCTTCTTCTCGAGCTGCTGTGGCCGTGGCCCTTCCTGTTCCAGAGCTGGGTAGG 455
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleCysSerTyrArgAspAspIleAspLeu 140
 DB 456 GACCGGTCCGGAGATTTCTTCAGAGCAACATCACTCTACCGGAGCATATGCTATG 515
 QY 141 GlnAsnLeuIleAspSerLeuGlnIlyAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160
 DB 516 CAAAACCTCATGACATCCCTTCAGAAAGCTAACCACTGCTGTGGCGCATATGGCCCTAA 575
 QY 161 AspTrpAspLeuAsnValIlyrPheAsnCysSerGlyAlaSerTyrSerArgGluAspCys 180
 DB 576 GACTGGGACCTCAACGCTCACTTCATTTGCAGCGGTGCCAGCTACAGCCGAGAGAAATGC 635
 QY 181 GlyValProPheSerCysCysValProAspProAlaGlnIlyValValAsnThrGlnCys 200
 DB 636 GGGGTCCCTTCTCTCTGCTCTGCTCGTCCAGATCTCGCCCAAAAGTTGTGAAACACAGTGT 695
 QY 201 GlyTyrAspValArgIleGlnLeuLysSerTyrAspGluSerIlePheThrIlyGly 220
 DB 696 GGATATGATGTACAGATTCAGTGAAGAGCAAGTGGATGAGTCCATCTTCACGAAAGGC 755
 QY 221 CysIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPhe 240
 DB 756 TGCATCCAGGGCGGTGGAACCTGGCTCCCGGGAACATTTTACATTTGTGGCTGGCGTTC 815
 QY 241 IleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAsp 260
 DB 816 ATCGCATCTGCTGTTGCAAGATATTGGCATCTTCTCGTCAAGACGCTGATCTCAGAC 875
 QY 261 IleGluAlaValIlyAlaGlyHisHisPhe 270
 DB 876 ATCGAGCGCATGGAAGACCGGCATCACTTC 905
 RESUR.T 11
 AA265350
 ID AA265350 standard; DNA; 1667 BP.
 XX
 XX AA265350;
 AC
 XX
 XX 23-MAR-2000 (first entry)
 DT
 XX
 DE Human secreted protein gene 10 fragment.
 KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy; chromosome 10; ds.
 KW
 XX Homo sapiens.
 OS
 XX
 XX W09958660-A1.
 PN
 XX 18-NOV-1999.
 PD
 XX
 XX 06-MAY-1999; 99WO-US09847.
 PF
 XX 12-MAY-1998; 98US-0085093.
 PR 12-MAY-1998; 98US-0085094.
 PR 12-MAY-1998; 98US-0085105.
 PR 12-MAY-1998; 98US-0085180.
 PR 18-MAY-1998; 98US-0085906.
 PR 18-MAY-1998; 98US-0085920.
 PR 18-MAY-1998; 98US-0085921.
 PR 18-MAY-1998; 98US-0085921.

PR 18-MAY-1998; 98US-0085922.
 PR 18-MAY-1998; 98US-0085923.
 PR 18-MAY-1998; 98US-0085924.
 PR 18-MAY-1998; 98US-0085928.
 PR 18-MAY-1998; 98US-0085925.
 PR 18-MAY-1998; 98US-0085927.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 PI Lafleur DM, Endress GA, Ebner R;
 XX WPI, 2000-062296/05.
 DR P-PSDB; AAY76266.
 XX
 PT New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Disclosure: Page 427-428; 475pp; English.
 XX
 CC AA265250 to AA265350 represent 97 isolated human secreted protein genes.
 CC This sequence was found to be present on human chromosome 10.
 CC AAY76124 to AAY76223 represent the secreted proteins encoded by the 97
 CC human genes. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 97 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
 CC secreted proteins.
 CC
 XX
 SQ Sequence 1667 BP; 323 A; 469 C; 488 G; 387 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 4.67e-147 Length: 1667
 Score: 1375.00 Matches: 267
 Percent Similarity: 98.53% Conservative: 1
 Best Local Similarity: 98.16% Mismatches: 2
 Query Match: 94.89% Indels: 3
 DB: 21 Gaps: 0
 US-09-905-674-2 (1-270) x AA265350 (1-1667)
 QY 1 MethISTYTYRARGTYSerASnAlaLysValSerCysTrpIlyrLysLeuIleuPhe 20
 DB 49 ATGCACATATATAGTACTTAACGCCAAGTCAAGCTGCTGACAGTACCTCTTTTC 108
 QY 21 SerTYRsnIleIlePheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAla 40
 DB 109 ACCTACACATCATCTCTGATGTGGCTGAGTTGCTCTTCTTGAGTGGGCGTGGGCA 168
 QY 41 TrpSerGlyLysGlyValLeuSerAspLeuThrLysValThrTrpMetHisGlyIleAsp 60
 DB 169 TGGACGCAAAAGGCTGCTCTCCGACCTCAACAAAGTGACCGGATGATGGAAATCGAC 228
 QY 61 ProValIleuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80
 DB 229 CCTGTGCTGTGCTCTCTGATGTGGCTGATGTTCACCTCGGGGCTGCGGCTGCTC 288
 QY 81 ValGlyValLeuArgGlyAsnIleCysLeuLeuAsnIlePheCysGlyThrIleValIleu 100
 DB 289 GTGGGGCTCTGCGGGAATATATCTGCTGCTCACTTTTCTGTGGACCACTGCTGCTC 348

QY 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArg 120
 DB 349 ATCTTCTTCTGGAGCTGCGTGGCCCTGCTGCTGCTCTCTGTTCCAGGACTGGGTGAGG 408
 QY 121 AspArgPheArgGluPhePheGluSerAsnIleLysSerTYRArgAspAspIleAspLeu 140
 DB 409 GACGGTTCGGGAGTCTCTTCGAGAGCAACATCAAGTCCATACGGGAGCATATCATCTCG 468
 QY 141 GlnAsnLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyValArgTyrProGlu 160
 DB 469 CAAMACCTATGACTCCCTTCAGAAACCTAACCAAGCTGTGGCCATATGGCCCTGAA 528
 QY 161 -AspTrpAspLeu-AsnValTYRAsnLysSerGlyValAspSerTYRArgGlyLysC 180
 DB 529 AGACTGGACCTCAGACGCTACTTCAATTTCACGGGTGCCAGCTACAGCCGAGAA-T 587
 QY 180 YsGlyValProPheSerCysCysValProAspProAlaGlnLysValValAsnThrGlnC 200
 DB 588 GCGGGGTCCCTTCTCTGCTGCTGCGCCAGATCCTGCGCAAAAAGTTGTGAAACACAGT 647
 QY 200 YsGlyTyrAspValArgIleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysG 220
 DB 648 GTGATATATATGCTCAGATTCAGCTGAGAGACAAAGTGATGATCATTTTCACAAAG 707
 QY 220 LysGlnIleGlnAlaLeuGluSerTrpLeuProArgAsnIleTYRLeuAlaGlyValP 240
 DB 708 GCTGCTCCAGCGCTGGAACCTGGCTCCCGCGAACAATTGATTTGCGCGGCTCT 767
 QY 240 IleIleAlaIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerA 260
 DB 768 TCAATGCCATTCGCTGCTGTTGAGATATTTGGCATCTTCTCGGCAAGACCTATCTCAG 827
 QY 260 sPTLeuAlaValLysAlaGlyHisHisPhe 270
 DB 828 ACATCGAGGACAGTGAAGCGCGCCATCACTTC 859
 RESULT 12
 ID AAA60182 standard; DNA; 639 BP.
 XX
 AC AAA60182;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Hydrophobic domain protein cDNA HP02956 isolated from KB cells.
 XX
 KW Human; secreted protein; membrane protein; hydrophobic domain;
 KW proliferation control; differentiation induction; material transport;
 KW biophysics; signal receptor; ion channel; transporter; immunostimulant;
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokine;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200029448-A2.
 XX
 PD 25-MAY-2000.
 XX
 PF 17-NOV-1999; 99WO-JP06412.
 XX
 PR 17-NOV-1998; 98JP-0326255.
 PR 22-DEC-1998; 98JP-0364315.
 PR 16-MAR-1999; 99JP-0069811.
 PR 27-APR-1999; 99JP-0119299.
 PR 19-MAY-1999; 99JP-0138169.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;

Db 310 GACTGGACCTCACGCTACTTCAATTCACAGCGGTGCCAGCTACAGCCGAGAGAAGTGC 366

channels and transporters. The present sequence is the coding sequence

CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.

XX Sequence 816 BP; 159 A; 230 C; 243 G; 184 T; 0 other;

Alignment Scores:

Pred. No.:	3,2e-110	Length:	816
Score:	1049.00	Matches:	204
Percent Similarity:	97.61%	Conservative:	0
Best Local Similarity:	97.61%	Mismatches:	4
Query Match:	72.39%	Indels:	2
DB:	21	Gaps:	0

US-09-905-674-2 (1-270) x AAA93675 (1-816)

```

QY 63 ValLeuValLeuMetValValMetPhehrLeuGlyPheAlaGlyCysValGly 82
   |||||
DB 5 GTCCTGCTCCTGATGGTGGCGCGGATGTCACCCG-GGGTTCGCGCGCGCTGGGG 63
   |||||
QY 83 Ala-LeuArgGluAsnIleCysLeuLeuAsnPhenheCysGlyThrIleValLeuIleph 102
   |||||
DB 64 GCGCGCCAGGAGAAATATCTGCTGCTCACTTTCTGTGGACCATCGTCATCTT 123
   |||||
QY 102 ePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAsp 122
   |||||
DB 124 CTCCTCGAGAGCTGCTGCTGGCGCTGCTGCTTCCTTCAGAGACTGGTGGAGACCG 183
   |||||
QY 122 gPheArgGluPhePheGluSerAsnIleCysSerTyraArgAspAspIleAspLeuGlnAs 142
   |||||
DB 184 GTTCCCGGGATTCTTCGAGAGACATCAAGTCTACCGCGGACGATGATCTCAAAA 243
   |||||
QY 142 nLeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTyraGlyProGluAspTr 162
   |||||
DB 244 CCTCATGACTCCTTCAAAAGCTTACCAAGCTGCTGTGGCATATGGCCTGAAAGACTG 303
   |||||
QY 162 PAspLeuAsnValTyraPheAsnCysSerGlyAlaSerTyraArgGluLysCysGlyVa 182
   |||||
DB 304 GAGACTCAAGCTACTTCAATGACGCGTCCAGCTACAGCCGAGAAAGTGGGGGT 363
   |||||
QY 182 lProPheSerCysValProAspProAlaGlnLysValValAsnThrGlnCysGlyTy 202
   |||||
DB 364 CCCCTTCTCTCTGCTGCTGCCAGATCTCCGCAAAAAGTTGGAACACACACTGTGGATA 423
   |||||
QY 202 rAspValArgGlnLeuLeuLysSerTyraPaspGluSerIlePheThrLysGlyCysI 222
   |||||
DB 424 TGAATGACAGATTCACTGAAGACAGAGTGAGTGAATCTTCAACGAAAGGCTGCAT 483
   |||||
QY 222 eGlnAlaLeuGluSerTrpLeuProArgAsnIleTyriLeValAlaGlyValPheIleAl 242
   |||||
DB 484 CCAAGGCGTGGAAAGCTGGCTCCCGGAAACATTTACATTTGGGCTGTCATCGC 543
   |||||
QY 242 aIleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleG 262
   |||||
DB 544 CATCTGCGTGTGGACATTTTGGCATTTCTCTGCGCAAGACGATGATCTCAGACATGA 603
   |||||
QY 262 uAlaValLysAlaGlyHisIsphe 270
   |||||
DB 604 GGCAGTGAAGGCCGCGCATCACTTC 628
   |||||

```

RESULT 15

AA880667 standard: cDNA: 601 BP.

AA880667;

13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16471.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

```

XX XX Homo sapiens.
OS OS
XX XX WO200175067-A2.
XX XX
XX XX 11-OCT-2001.
XX XX
XX XX 30-MAR-2001; 2001WO-US08631.
XX XX
XX XX 31-MAR-2000; 2000US-0540217.
XX XX 23-AUG-2000; 2000US-0549167.
XX XX
XX XX (HYSE-) HYSEQ INC.
XX XX
XX XX Dermanac RT, Liu C, Tang YT;
XX XX WPI, 2001-639362/73.
XX XX P-PSDB: ABG16480.
XX XX
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX XX diagnostics, forensics, gene mapping, identification of mutations
XX XX responsible for genetic disorders or other traits and to assess
XX XX biodiversity.
XX XX
XX XX Claim 1; SEQ ID No 16471; 103pp; English.
XX XX
XX XX The invention relates to isolated polynucleotide (I) and
XX XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX XX and gene mapping, and in recombinant production of (II). The
XX XX polynucleotides are also used in diagnostics as expressed sequence tags
XX XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX XX to restore normal activity of (II) or to treat disease states involving
XX XX (II). (II) is useful for generating antibodies against it, detecting or
XX XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX XX a food supplement. (II) and its binding partners are useful in medical
XX XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX XX disorders involving aberrant protein expression or biological activity.
XX XX The polypeptide and polynucleotide sequences have applications in
XX XX diagnostics, forensics, gene mapping, identification of mutations
XX XX responsible for genetic disorders or other traits to assess biodiversity
XX XX and to produce other types of data and products dependent on DNA and
XX XX amino acid sequences. AA880667-AAS94564 represent novel human
XX XX diagnostic coding sequences of the invention.
XX XX Note: The sequence data for this patent did not appear in the printed
XX XX specification, but was obtained in electronic format directly from WIPO
XX XX at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX XX Sequence 601 BP; 107 A; 168 C; 180 G; 146 T; 0 other;
XX XX
XX XX Alignment Scores:
XX XX Pred. No.: 1,72e-97 Length: 601
XX XX Score: 936.00 Matches: 175
XX XX Percent Similarity: 98.31% Conservative: 0
XX XX Best Local Similarity: 98.31% Mismatches: 3
XX XX Query Match: 64.60% Indels: 0
XX XX DB: 23 Gaps: 0
XX XX
XX XX US-09-905-674-2 (1-270) x AA880667 (1-601)
XX XX
QY 1 MethIstYrYrArgYrSerAsnAlaLysValSerCysTrpYrIstYrLeuLeuPhe 20
   |||
DB 67 ATGCGTATGATAGATACTTAACGCCAAGTCAAGCTGTGTACAGTACCTCTTTC 126
   |||
QY 21 SerTyraSnIleIlePheTrpLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAla 40
   |||
DB 127 AGCTAACACATCATCTTCTGCTGGTGGAGTGTCTTCTTGAGAGTGGGCTGTGGCA 186
   |||
QY 41 TrpSerGluLysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAsp 60
   |||
DB 187 TGGAGCCAAAGGGGTGTGCTCTCCGACCTACCAATGACCCGAGATGGAATGAC 246
   |||
QY 61 ProValValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCys 80
   |||

```

```

Db 247 COTGGTGTCTGTCTGATGTGGGGGTGATGTTACCCCTGGGGGTTCGCCGGCTGC 306
QY 81 ValGlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeu 100
Db 307 GTGGGGGCTTCGCGGAGAAATCTGCTCAACTTTTCTGTGTGGCACCACATCGTCTC 366
QY 101 IlePhePheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGluAspTrpValArg 120
Db 367 ATCTTCTTCTCTGAGAGCTGTGTGGCCGAGCTGTGCTCTCTGTTCCAGGACTGGGTGAGG 426
QY 121 AsparGpheaArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeu 140
Db 427 GACCGGTTCGCGGAGTTCTTCGAGAGCAACATCAAGTCTTACCGGCGAGATATGATCTG 486
QY 141 GluAsnLeuIleAspSerLeuGluLysAlaAsnGlnCysCysGlyAlaTyrGlyProGlu 160
Db 487 CAAAACCTCATCTGACTCCCTTCAGAAAGCTTACCAGTGTGTGGCGCATATGGCCCTGAA 546
QY 161 AspTrpAspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlu 178
Db 547 GACTGGGAGACTCAAGCTTACTCAATTGCAGCGGTGCCAGCTACAGCCGAGAG 600

```

Search completed: December 19, 2002, 06:26:57
 Job time : 269 secs

	LOCATION (1564)	n equals a,t,g, or c	
	OTHER INFORMATION: n equals a,t,g, or c		
	NAME/KEY: misc_feature		
	LOCATION: (1623)		
	OTHER INFORMATION: n equals a,t,g, or c		
	NAME/KEY: misc_feature		
	LOCATION: (1643)		
	OTHER INFORMATION: n equals a,t,g, or c		
US-09-925-299-205			
Alignment Scores:			
Pred. No.:	1,68e-95	Length:	1655
Score:	867.00	Matches:	154
Percent Similarity:	76.25%	Conservative:	45
Best Local Similarity:	59.00%	Mismatches:	62
Query Match:	59.83%	Indels:	0
DB:	10	Gaps:	0
US-09-905-674-2 (1-270) x US-09-925-299-205 (1-1655)			
OY	6 TyISerAsnAlaValSerCysTrpIlyrIyLeuLephSeSerTyRAsnIlle	25	
Db	292 TACAGGGCTCCGTGAAGTCATGTTCCATCAACAATCTTCAATTGGCGTTAAAGTATA	351	
OY	26 PheTrpLeuAlaGlyValAlaPheLeuGlyAlaGlyLeuTrpAlaTrpSerGluLySgLy	45	
Db	352 TTGTGGTTTTGGAAATTAACATTCTTGGAATTTGGACTGTGGCATGGATGAAGAAGA	411	
OY	46 ValLeuSerAspLeuThrIlyValThrArgMetHisGlyIleAspProValIleuVal	65	
Db	412 GTTCTGTCCAACATCTTCCATCACACCGAATTCGGGGCTTGACCAGCTTTGGCTGTC	471	
OY	66 LeuLeuValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArg	85	
Db	472 CTGTGGGGGAGGAGGATGATTCATTTTGGGATTTGCAGGGTGCAATGGAGCGCTACGG	531	
OY	86 GluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhePheLeuGlu	105	
Db	532 GAAAACACTTTCCTTCCAAGTTTCTGTCTCTTCTGGAAATTAATTTCTCTGGAG	591	
OY	106 LeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValaArgAspArgPheArgLu	125	
Db	592 CTCACCTGCCGAGTTCTTAGCATTTGTTTCAAGACTGGATCAAGACACGCTGATATTC	651	
OY	126 PhePheGluSerAsnIleLySserTyRArgAspAspIleAspLeuGlnAsnLeuIleAsp	145	
Db	652 TTATTAACAACAACATCAAGCATATCCGGATGACATTTGATTCAAAACTCATATAC	711	
OY	146 SerLeuGlnLyAlaAsnGlnCysGlyAlaTyGlyProGlnAlaSpTrpAspPheAsn	165	
Db	712 TTCACCCGGAATTTGGCGTGTGGGGCTTTGGAGCTGATCATTTGGAAUCCATAAT	771	
OY	166 ValTyRPhenAsnCysSerGlyAlaSerTySerArgIlyLysCysGlyValProPheSer	185	
Db	772 ATTACTTCAATTGCAAGATTCCAATGCAAGTCGAGAACCATGTGGCTTCATTCC	831	
OY	186 CysCysValProAspProAlaGlnLyValAlaAsnThrGlnCysGlyTyRAspValaArg	205	
Db	832 TGCTGCACATAAAGATCCCGGAGGAAGATGTCATCAACATCAGTGHGTATATATCCAGG	891	
OY	206 IleGlnLeuLySerLySTrpAspGluSerIlePheThrIlySgLyIleGlnAlaLeu	225	
Db	892 CAAMAACCGAAGTTGACCAACAGATTTGATCTTACACGAAGGCTGTGCCACAGTTT	951	
OY	226 GluSerTrpLeuProArgAsnIleTyRIleValAlaGlyValaPheIleAlaIleSerIeu	245	
Db	952 GAGAAAGTGTTGCAGAGCAATTTAACCATGTTGCTGTGATTTTCAATAGCATTTCATTG	1011	
OY	246 IeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGluAlaValLyS	265	
Db	1012 CTGCAGATATTTGGGATATGCTGTGCCAGAAATTTGGTATGCAATATCAAGCTGTACAGG	1071	
OY	266 Ala 266		

```

Db      1072 GCG 1074
||
RESULT 2
US-09-934-268-3
; Sequence 3, Application US/09934268
; Patent No. US20020172986A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: 23228, A NOVEL HUMAN TETRASPANIN FAMILY
; TITLE OF INVENTION: MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-079001
; CURRENT APPLICATION NUMBER: US/09/934, 268
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226, 612
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-934-268-3

Alignment Scores:
Pred. No.:      7,15e-95      Length:      813
Score:          858.00      Matches:      154
Percent Similarity: 75.76%      Conservative: 45
Best Local Similarity: 58.33%      Mismatches:  64
Query Match:    59.21%      Indels:      0
DB:            9      Gaps:      0

US-09-905-674-2 (1-270) x US-09-934-268-3 (1-813)
*
Qy      3 TyrTyrAargTyrSerAsnAlaIysValSerCysTrpTyrIlystYrLeuLeuPheSerTyr 22
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      13 CACCAAGCATTTCCAGAGAACCTCGAGGTGGCGTGGCGGGAATAACTCTCTCGTTGGCTTC 72
Qy      23 AsnIleIlePheTrpLeuAlaIglValValPheLeuValAlaIglYleuTrpAlaTrpSer 42
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      73 AACATTTGCTTCTGGGTCTGGAGCCCTGTCTCGCTTCGCGCTTCGCGCTTCGGCTGGGGT 132
Qy      43 GluTysGlyValLeuSerAspLeuThrLysValThrArgMetHisGlyIleAspProVal 62
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      133 GAGAAAGGCGCTTCTCTCGAACATCTCAGCGGTGACAGATCTGGGAGAGGCCCTTGACCCGTG 192
Qy      63 ValLeuValLeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGly 82
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      193 TGGCTGTGTGTGGTGTGGTGGAGCGGTATGTGCGTGGCTTGTGGCTTCGATGGG 252
Qy      83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValLeuIlePhe 102
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      253 GCCCTCCGGGAGAACACCTTCTCTCTCAAGTTTCTCCGTGTTCCTCGGCTTCATCTTC 312
Qy      103 PheLeuGluLeuAlaValAlaValLeuAlaPheLeuPheGlnAspTrpValArgAspArg 122
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      313 TTCCTGGAGCTGGCACAAGGAGTCTGGCTTGTCTTCAAGAGACTGGATTGAGAACAG 372
Qy      123 PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn 142
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      373 CTCACCTTCTCATCAACACACAGCTCAAGGCTCACCGGGGACGACATGTGACTCCAGAAC 432
Qy      143 LeuIleAspSerLeuGlnLysAlaAsnGlnCysGlyAlaTrpGlyProGluAspTrp 162
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      433 CTCATGATCTTGTGCTCAGAGATCTGCTTGTGTGGAGCCCGAGGCCCCCAATGACTGG 492
Qy      163 AspLeuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGluTysCysGlyVal 182
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      493 AACCTCAATATCTACTTCACTGACACTACTTGAACCCCAAGCCCGGAGCCCTCGGGGTG 552
Qy      183 ProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyr 202
      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      553 CCTTCTCTCTGTGGTGGTGAAGACCCTCGGAGAGATGTCTCAACACCAAGTGTGCTAC 612

```


Qy	43	GlulysgylValleuSerAspleuthrLysValThrArgmetHisgylIleasProval	62
Db	251	GAGAAAGGGTGTCTCTCCAACTCTCCGCGTGAACCCATCTGGAGGCCCTGCACCTGTG	310
Qy	63	ValleuValleuMetValgIValValMetPhehrhLeuglyPhealagIcysValgIy	82
Db	311	TGGCTGTTTGAAGTGTGGAGGGCTCATGTGCCGTGGCTTGGCGGCTGCATGGG	370
Qy	83	AlaIeuArgIuAsnIleCysLleuIleAsnPhetheCysgylThrIleValleuIlePhe	102
Db	371	GCTCTCGGGAGAACACTTTCCTGCTCAAGTTTTTCAGTTCCTTGGCCATCTTC	430
Qy	103	PheIeuGluLeuAlaValAlaValLeuAlaPheIeuPheGlnsptRyValArgaspArg	122
Db	431	TTCTCGAGCGTGGACACAGGATCTTGGCTTCCTATTCAAAGCACTGGATTCGAGCACG	490
Qy	123	PheArgGluPhePheGluSerAsnIleLysSerThrArgaspIleaspleuGlnAsn	142
Db	491	CTCATTTCTTCATTAAACACACAGCTCAAGGCTATCGGATGACATTGACCTCCAGAAC	550
Qy	143	LleuIleasPserLeuGlnLysAlaAsnGlnCysCysgylAlaIArgIlyProGluAspTrp	162
Db	551	CTCATTTGACTTTCGCTCAGAAATATTGGTCTGCTGGAGACCCGAGGGCTAATGACTGG	610
Qy	163	AspleuAsnValIYrPheAsnCysSerGlyValAserTYrSerArgIuLysCysgylVal	182
Db	611	AACCTCATATCTTCTTTCACATGCACGACTTGACACCCGAGCGAGCGCTCGGGGGTG	670
Qy	183	ProPheSerCysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysgylYr	202
Db	671	CCCTTCTCTCTCTGTGTGCAGGGACCTGCGCAT---GTCTCAACACCCAGTGTGGCTAT	726
Qy	203	AspValArgIleGlnLysSerLysTYrAspIuSerIlePheThrLysgIlyCysIle	222
Db	727	GATGTCCGGCTCAACAGTGAAGCTGACAGACAGAGGCTCCATACACCAAGGCTGTGTG	786
Qy	223	GlnIleAlaLeuGluSerTrpLeuProArgAsnIleYrIleValAlaGlyValAlaPheIleAla	242
Db	787	GGCCAGTTTGGAAGTGGCTGCACAGACCACTGATCGTGTGGTGGGGTCTTTGGTGGCC	846
Qy	243	IleSerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlu	262
Db	847	ATCGCTCTCCCTCGAAGCTTTGGTATGTGCTGGCCCAAGAACTGTGATGATCATCAAG	906
Qy	263	AlaValAlysAla 266	
Db	907	GCAGTGAAGGCC 918	
RESULT 5			
US-09-729-674-169			
Sequence 169, Application US/09729674			
Patent No. US2001003935A1			
GENERAL INFORMATION:			
APPLICANT: Jacobs, Kenneth			
APPLICANT: McCoy, John M.			
APPLICANT: LaValle, Edward R.			
APPLICANT: Collins-Racie, Lisa A.			
APPLICANT: Evans, Cheryl			
APPLICANT: Merberg, David			
APPLICANT: Treacy, Maurice			
APPLICANT: Agostino, Michael J.			
APPLICANT: Steindinger II, Robert J.			
APPLICANT: Spaulding, Vikki			
APPLICANT: Wong, Gordon G.			
APPLICANT: Clark, Hillary			
APPLICANT: Fechtel, Kim			
APPLICANT: Genetics Institute, Inc.			
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM			
FILE REFERENCE: 6055-64X			
CURRENT APPLICATION NUMBER: US/09/729,674			
CURRENT FILING DATE: 2000-12-04			
PRIOR APPLICATION NUMBER: 09/539,330			

	:	PRIOR FILING DATE:	2000-03-30
	:	NUMBER OF SEQ ID NOS:	283
	:	SOFTWARE:	Patentin Ver. 2.0
	:	SEQ ID NO	169
	:	LENGTH:	1110
	:	TYPE:	DNA
	:	ORGANISM:	Homo sapiens
	:	US-09-729-674-169	
		Alignment Scores:	
		Pred. No.:	4,35e-79
		Score:	729.50
		Percent Similarity:	68.20%
		Best Local Similarity:	53.64%
		Query Match:	50.35%
		DB:	10
			Gaps: 2
		US-09-905-674-2 (-1-270) x US-09-729-674-169 (-1-1110)	
OY	6	TyrSerAsnAlaIysValSerCysTyrPylrLysThrLeuLeuPheSerTyrrAsnIlelle	25
Db	191	TACAAAGGCTCCTGAAGTAGTGTTGCATCAAAACTTCATTGGCTCCATATGTATA	250
OY	26	PheTrpLeuAlaGlyValValPheLeuGlyValGlyLeuTrpAlaTrpSergLutLySgly	45
Db	251	TTTTGCGTTTGGGAATACATTTCTGGAGTTGGACGTGGCGCATGGATGAAAAAGA	310
OY	46	ValLeuSerAspLeuThrlLysValThrArgMetHisglYlLeasProvalValLeuVal	65
Db	311	GTTCGTGCCAACATCTCTTCATCACCGATCTCGGCGGCTTTGACCAGTTTGGCTCTTC	370
OY	66	LeuMetValGlyValValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArg	85
Db	371	CTTGtGGtGGGAGGAGtGAtGTTcATTtTGGgATTtTGGAgGtTCATtTGAGCGctTAAG	430
OY	86	GluAsnIleCysLeuLeuAsnPhePheCysGlYThrIleValLeuIlePhePheLeuGlu	105
Db	431	GA AAAA CACTTCTCTCAAGCTTTTCTCTGTCCTCGGAATATTATTTCTTCCTCGAG	490
OY	106	LeuAlaValAlaValLeuAlaPheLeuPheGlnSprTrpValArGAspArgPheArgGlu	125
Db	491	CTCActGCCGAGtTTCtAGcAttTtGtTTtCAAAGAcTgcGATcCAAGAGccAGctGTATtTC	550
OY	126	PhePheGluSerAsnIleLysSerTyrTrArgAspAlleasPheLeuGlnAsnLeuIleasp	145
Db	551	TTTATTAACAACAACATCAGACATATCGGGATGACATGTATTGCAAAACCTCATAGAC	610
OY	146	SerLeuGlnLysAlaAsnGlnCysCysGlyAlaIryelYProGluAsprTPAspLeuAsn	165
Db	611	TTTCAACCCGGA-----	621
OY	166	ValTyrPheAsnCysSergLYlaSerTyrSerArgLutLysCysGlyAlaProPheSer	185
Db	622	ATATATTCATGCA-----AGTGAAGAGGAGTGTGGCGTTCCATTCTCC	666
OY	186	CysCysValProAspProAlaGlnLysValAlaAsnThrGlnCysGlyTyrAspValArg	205
Db	667	TGCTGCACATAAAGATCCCGCGAGAAGATGCATCAACACTCAGTGTGCTATGATGCCAGG	726
OY	206	IleGlnLeuLysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeu	225
Db	727	CAAAAACCAAGAGTTGCACAGCAGATTGTATCTACACGAAAGGCTGTGCCCCCACTTT	786
OY	226	GluSerTrpLeuProArgAsnIleTyrIleValAlaGlyAlaPheIleAlaIleSerLeu	245
Db	787	GAGAAGTGTTGCAGAGCAATTAACCMCTGGCTGTATTTTCAATAGCATTGCATTG	846
OY	246	LeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerSprIleGlnAlaValLys	265
Db	847	CTGCAGATATTTGGGATTTGGCTGGCCCAAAATTTGGTATGCGATATCGAGCTGTACG	906
OY	266	Ala 266	

Db 907 GCG 909

RESULT 6

US-09-875-440-1

; Sequence 1, Application US/09875440

; Patent No. US20020156035A1

; GENERAL INFORMATION:

; APPLICANT: Reinhard, Christoph

; APPLICANT: Jefferson, Anne B.

; APPLICANT: Winter, Jill A.

; APPLICANT: Randazzo, Filippo

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

; TITLE OF INVENTION: NEOPLASTIC DISEASE USING NET-4 MODULATORS

; FILE REFERENCE: PP-01701.002/200130.522

; CURRENT APPLICATION NUMBER: US/09/875.440

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 864

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: NET-4 oligonucleotide used in cell proliferation

; OTHER INFORMATION: assay on SW620 cells

US-09-875-440-1

Alignment Scores:

Pred. No.:	1.51e-59	Length:	864
Score:	567.00	Matches:	98
Percent Similarity:	77.11%	Conservative:	30
Best Local Similarity:	59.04%	Mismatches:	38
Query Match:	39.13%	Indels:	0
DB:	9	Gaps:	0

US-09-905-674-2 (1-270) x US-09-875-440-1 (1-864)

QY 6 TYSERASNAALYSAVALSERCYSTRPYRILSYTYLEULEUPHESERTYRASNILLE 25

DB 363 TPCAAAGGCGCCGAGAGTCAAGTTGTCATCAATCTTCATATTTGGCTTAAGTCA 422

QY 26 PHERPLLEUALGLYVALVALPHELEUGLYVALGLYLEUTRPAATRPSERGLULYSGLY 45

DB 423 TTTTGTTTGGTAATACATTTCTTGGAATGTGACGTGGGCAATGAAGAAAAAGA 482

QY 46 VALLEUSERASPLEUTHLYSVALTHARGMETHISGLYILEASPRIVALLLEUVAL 65

DB 483 GTTCGTCCAAACATCTCTCCATCCACCATCTCGGCGGCTTGACCCAGTTGGCTCTTC 542

QY 66 LEUMETVALGLYVALVALMERPHERTHRARGMETHISGLYILEASPRIVALLLEUVAL 85

DB 543 CTTGGTGCGGAGGAGGAGTGTTCATTTGGGATTTGAGGGGTGCATTGGAGCGCTACGG 602

QY 86 GLUASNILECYSALEULEUASNPHENECYSGLYTHRILEVALLEULEUPHLEULEU 105

DB 603 GAAACACACTTCTCTCAAGTTTCTGTCGTCCGGAATATTTCTCTCTGAG 662

QY 106 LEUALVALVALLEUALAPHELEUPHEGLINASPIRVALLANGASPARPHEARGLU 125

DB 663 CTCACATGCCGAGTCTAGCATTTGTTTCAAGACGTGATCAAGACACGCTGATATTC 722

QY 126 PHEPHEGLUSERASNILEYSERTYRARGASPARSPLLEASPLEUGLNASNLEULEASP 145

DB 723 TTTATTAACACACATCAAGCATATGAGCATGACATGATGCAAAACCTCATTAAC 782

QY 146 SERLEUGLNYALAAAGLNGLNCYSCYSELALATYRGLYPROGLUASPTRPASPLEUASN 165

DB 783 TTCACCCAGGAAATATTGGAGCGCTGTGGGCTTTTGGAGCTGATGATTTGAACCTAAT 842

QY 166 VALTYRPHENARGYSSER 171

DB 843 ATTACTCATCAATTGCACA 860

RESULT 7

US-10-052-586-277

; Sequence 277, Application US/10052586

; Patent No. US20020127584A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gunney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C1

; CURRENT APPLICATION NUMBER: US/10/052.586

; CURRENT FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063486

; PRIOR FILING DATE: 1997-10-21

; PRIOR APPLICATION NUMBER: 60/063540

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063541

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063544

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063564

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/063734

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/063870

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/064103

; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066120

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/066466

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/066772

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/069335

; PRIOR FILING DATE: 1997-12-11

; PRIOR APPLICATION NUMBER: 60/069425

; PRIOR FILING DATE: 1997-12-12

; PRIOR APPLICATION NUMBER: 60/069870

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/068017

; PRIOR FILING DATE: 1997-12-18

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/078886

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078939

; PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079366	PRIOR APPLICATION NUMBER: 60/083322	PRIOR APPLICATION NUMBER: 60/083495	PRIOR APPLICATION NUMBER: 60/083585	PRIOR APPLICATION NUMBER: 60/083666	PRIOR APPLICATION NUMBER: 60/084141	PRIOR APPLICATION NUMBER: 60/084639	PRIOR APPLICATION NUMBER: 60/084640	PRIOR APPLICATION NUMBER: 60/084643	PRIOR APPLICATION NUMBER: 60/085077	PRIOR APPLICATION NUMBER: 60/085573	PRIOR APPLICATION NUMBER: 60/085797	PRIOR APPLICATION NUMBER: 60/085815	PRIOR APPLICATION NUMBER: 60/085850	PRIOR APPLICATION NUMBER: 60/085852	PRIOR APPLICATION NUMBER: 60/085855	PRIOR APPLICATION NUMBER: 60/085700	PRIOR APPLICATION NUMBER: 60/086023	PRIOR APPLICATION NUMBER: 60/086028	PRIOR APPLICATION NUMBER: 60/086392	PRIOR APPLICATION NUMBER: 60/086466	PRIOR APPLICATION NUMBER: 60/086609	PRIOR APPLICATION NUMBER: 60/087599	PRIOR FILING DATE: 1998-06-02	PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-08	PRIOR FILING DATE: 1998-04-29	PRIOR FILING DATE: 1998-04-29	PRIOR FILING DATE: 1998-05-05	PRIOR FILING DATE: 1998-05-06	PRIOR FILING DATE: 1998-05-07	PRIOR FILING DATE: 1998-05-07	PRIOR FILING DATE: 1998-05-07	PRIOR FILING DATE: 1998-05-07	PRIOR FILING DATE: 1998-05-15	PRIOR FILING DATE: 1998-05-15	PRIOR FILING DATE: 1998-05-15	PRIOR FILING DATE: 1998-05-18	PRIOR FILING DATE: 1998-05-22	PRIOR FILING DATE: 1998-05-22	PRIOR FILING DATE: 1998-05-22	PRIOR FILING DATE: 1998-05-28	PRIOR FILING DATE: 1998-05-28	PRIOR FILING DATE: 1998-05-28	PRIOR FILING DATE: 1998-05-28	PRIOR FILING DATE: 1998-06-02	PRIOR FILING DATE: 1998-06-02	PRIOR FILING DATE: 1998-06-02	

1	PRIOR FILING DATE: 1998-06-03
2	PRIOR APPLICATION NUMBER: 60/088023
3	PRIOR FILING DATE: 1998-06-04
4	PRIOR APPLICATION NUMBER: 60/088028
5	PRIOR FILING DATE: 1998-06-04
6	PRIOR APPLICATION NUMBER: 60/088025
7	PRIOR FILING DATE: 1998-06-04
8	PRIOR APPLICATION NUMBER: 60/088033
9	PRIOR FILING DATE: 1998-06-04
10	PRIOR APPLICATION NUMBER: 60/088167
11	PRIOR FILING DATE: 1998-06-05
12	PRIOR APPLICATION NUMBER: 60/088202
13	PRIOR FILING DATE: 1998-06-05
14	PRIOR APPLICATION NUMBER: 60/088212
15	PRIOR FILING DATE: 1998-06-05
16	PRIOR APPLICATION NUMBER: 60/088217
17	PRIOR FILING DATE: 1998-06-05
18	PRIOR APPLICATION NUMBER: 60/088326
19	PRIOR FILING DATE: 1998-06-04
20	PRIOR APPLICATION NUMBER: 60/088555
21	PRIOR FILING DATE: 1998-06-09
22	PRIOR APPLICATION NUMBER: 60/088722
23	PRIOR FILING DATE: 1998-06-10
24	PRIOR APPLICATION NUMBER: 60/088738
25	PRIOR FILING DATE: 1998-06-10
26	PRIOR APPLICATION NUMBER: 60/088740
27	PRIOR FILING DATE: 1998-06-10
28	PRIOR APPLICATION NUMBER: 60/088811
29	PRIOR FILING DATE: 1998-06-10
30	PRIOR APPLICATION NUMBER: 60/088824
31	PRIOR FILING DATE: 1998-06-10
32	PRIOR APPLICATION NUMBER: 60/088825
33	PRIOR FILING DATE: 1998-06-10
34	PRIOR APPLICATION NUMBER: 60/088826
35	PRIOR FILING DATE: 1998-06-10
36	PRIOR APPLICATION NUMBER: 60/088861
37	PRIOR FILING DATE: 1998-06-11
38	PRIOR APPLICATION NUMBER: 60/088863
39	PRIOR FILING DATE: 1998-06-11
40	PRIOR APPLICATION NUMBER: 60/088876
41	PRIOR FILING DATE: 1998-06-11
42	PRIOR APPLICATION NUMBER: 60/088909
43	PRIOR FILING DATE: 1998-06-12
44	PRIOR APPLICATION NUMBER: 60/089105
45	PRIOR FILING DATE: 1998-06-12
46	PRIOR APPLICATION NUMBER: 60/089512
47	PRIOR FILING DATE: 1998-06-16
48	PRIOR APPLICATION NUMBER: 60/089514
49	PRIOR FILING DATE: 1998-06-16
50	PRIOR APPLICATION NUMBER: 60/089538
51	PRIOR FILING DATE: 1998-06-17
52	PRIOR APPLICATION NUMBER: 60/089589
53	PRIOR FILING DATE: 1998-06-17
54	PRIOR APPLICATION NUMBER: 60/089653
55	PRIOR FILING DATE: 1998-06-17
56	PRIOR APPLICATION NUMBER: 60/089658
57	PRIOR FILING DATE: 1998-06-17
58	PRIOR APPLICATION NUMBER: 60/089908

Alignment Scores:	
Pred. NO.:	2,78e-38
Score:	395.50
Percent Similarity:	32.56
Best Local Similarity:	31.758
Query Match:	27.29%
DB:	12
Gaps:	3
Length:	1778
Matches:	81
Conservative:	53
Mismatches:	114
Indels:	7
Gaps:	3

US-09-905-674-2 (1-270) x US-10-052-586-277 (1-1778)

[illegible]

```
Db 309 CTGCTGTGGGCATCTATGACAGAGTTGAGCGCGAGAAATATTAACCCCTTGAAAGTCC 368
Qy 54 ThrArgMetHisGlyIleAspProValIleuValIleuMetValGlyValIleuMetPhe 73
Db 369 TTC-----CTGGCTCCAGCCATCATCCATCCCTCTGGCGGTGCTCATGTTTC 416
Qy 74 ThrLeuGlyPheAlaGlyCysValAlaIleuArgIleuAsnIleCysLeuLeuAsnPhe 93
Db 417 ATGGTCTCCTCATGTTGGTGTGGTGGTCCCTCCGTGACAACTGTACTCTCTCCAAAGC 476
Qy 94 PheCysGlyThrIleValIleuLeuPhePheLeuGluIleuAlaValAlaValIleuAlaPhe 113
Db 477 TTCATGATACATCCCTGGAGATGCTCCCTCATCATGAGAGCATGTTGGTGGTGGTCC 536
Qy 114 LeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIleLeuSer 133
Db 537 ACCTTCCGGAACACACCATTTGACTTCTGAAACGACAACTTCGAAAGAGGATTCAGAAC 596
Qy 134 TyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnIleValAlaAsnGlnCys 153
Db 597 TACTATGATGATCTGTGACTTCAAAAACATCATGAGACTTGTTCAGAAAAAGTTCAGTGC 656
Qy 154 CysGlyAlaIleTyrGlyProGlnAspTrpAspLeuAsnValTyrPheAsnCysSerGlyAla 173
Db 657 TGTGGCGGGGAGACTACCGAGATTGAGAGCAAGAAATCAGTACACAGACTGCAGTCCCT 716
Qy 174 SerTyrSerArgGluLysCysGlyValProPheSerCysValProAspProAlaGln 193
Db 717 -----GGACCCCTGGCTGTGGGTGGCTGACCTGCTCATGAGAGACGACACAA 770
Qy 194 LysValIleAsnThrGlnCysGlyTyrAspValArgIleGlnIleuLysSerIleTrpAsp 213
Db 771 ---GTTGTCACACCATGATGTTGGTGCACAAACTATCGACAAAGACGTTTCAGTGTGAG 827
Qy 214 GluSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrpLeuProArgAsnIle 233
Db 828 GATGATCATCTACGTGGGGCTGCACCAACGCCGTGATCATCTGTTGATGACAACTAC 887
Qy 234 TyrIleValAlaGlyValPheIleAlaIleSerIleuLeuGlnIlePheGlyIlePheLeu 253
Db 888 ACCATCATGCGCTGATCTCTCTGGGCACTCTGCTTCCCGAATTCCTGGGGGTGCTGCTG 947
Qy 254 AlaArgThrIleuLeuSerAspIleGluAlaValIleuLysAlaGlyHis 268
Db 948 AGCGTGTGCTGATCATCACCCGGGTGAGAGACATCATCATGAGAGCAC 992

RESULT 8
; Sequence 279, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Devin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Filing, Steve P.
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864, 864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 279
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-864-864-279
Alignment Scores:
Pred. No.: 3,16e-27 Length: 873
Score: 300.50 Matches: 72
Percent Similarity: 48.41% Conservative: 50
Best Local Similarity: 28.57% Mismatches: 97
Query Match: 20.74% Indels: 33
DB: 10 Gaps: 6

US-09-905-674-2 (1-270) x US-09-864-864-279 (1-873)
Qy 16 LysTyrIleuLeuPheSerTyrAsnIleIlePheTrpLeuAlaGlyValIleuLeuGly 35
Db 100 AAGTTTGTGCTACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 159
Qy 36 ValGlyLeuTrpAlaTrpSerGlyLysGlyValIleuSerAspLeuThrLysValThrArg 55
Db 160 GTGGGTGCTGGGGCA-----CAGCTTGTCTGTGAGTCA-----ACCATTA 198
Qy 56 MetHisGlyIleAspPro-----ValIleuValIleuMetValGlyValIleu 71
Db 199 ATCCAGGGGGCTACCCCTGCTCTGTTGCCAGTGTGTCATTCGACGAGTGTGCTTTC 258
Qy 72 MetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeuLeu 91
Db 259 CTCTTCTGATGGCTTTTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
Qy 92 AsnPheCysGlyThrIleValIleuLeuIlePhePheLeuGluLeuAlaValAlaIleu 111
Db 319 ATCAGCTTTGCTCATCTTCTGCTCTTATCATGTTGTTGAGTGGGCGCCGACCATGCT 378
Qy 112 AlaIleuLeuPheGlnAspTrpValAlaArgAspArgPheArgGluPhePheGluSerAsnIle 131
Db 379 GCGATGTGTTTACAGTAAGTGAATGTCAGAGTTATTAACAATTCGCGCAGCAGATG 438
Qy 132 LysSerTyrArgAspAspIleAspLeuGlnAsnLeuIleAspSerLeuGlnIleValAlaAsn 151
Db 439 GAGATTAACCCGAAACAAACACACACCTTCATCTGTCAGACAGATGACGACGATTTT 498
Qy 152 GlnCysCysGlyAlaIleTyrGlyProGlnAspTrpAspLeuAsnValTyrPheAsnCysSer 171
Db 499 AAGTGTGTGGGGCTGCTACTACACAGATTGGAG----- 534
Qy 172 GlyIleAspTrpSerArgGluLysCysGlyValProPheSerCysValProAspPro 191
Db 535 ---AAAAATCCCTTCATGTCGAAAGAACCGAGTCCCGACCTCGCTGCATTAAT----- 585
Qy 192 AlaGlnLysValIleAsnThrGlnCysGlyTyrAspValArgIleGlnIleuLysSerIle 211
Db 586 -----GTTACTGTGGGCTGTGGGATTAATTTCAC----- 615
Qy 212 TrpAspLeuSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrpLeuProArg 231
Db 616 ---CAGAAGCGATCATTAAGAGAGGCTGTGTGAGAAATGCGGGGCTGCTGAGAGAAA 672
Qy 232 AsnIleTyrIleValAlaGlyValPheIleAlaIleSerIleuLeuGlnIlePheGlyIle 251
Db 673 AATGTGCTGTGTGAGTGTGAGCAGCGCCCTGGAATGCTTTTGTGAGGTTTGGGAAT 732
Qy 252 PheLeuAlaArgThrIleuLeuSerAspIleGluAla 263
Db 733 GTCTTTCCTGCTGCTGCTGCTGAGAGATATCAGAAGT 768

RESULT 9
; Sequence 31, Application US/09823356
; Patent No. US20010025098A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
```



```

; NAME/KEY: misc feature
; LOCATION: (400)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (408)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (411)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (415)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (436)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10353

Alignment Scores:
Pred. No.: 1.53e-25 Length: 446
Score: 283.00 Matches: 60
Percent Similarity: 95.31% Conservative: 1
Best Local Similarity: 93.75% Mismatches: 1
Query Match: 19.53% Indels: 2
DB: 10 Gaps: 0

US-09-905-674-2 (1-270) x US-09-783-590-10353 (1-446)
QY 209 LysSerLysTrpAspGluSerIlePheThrLysGlyCysIleGlnAlaLeuGluSerTrp 228
DB 6 AGAAGCAAGTGGATGAGTCCATCTTCACGAAAGCGCTGATCGAGCGCTGGAAGCTGG 65
QY 229 LeuProArg-AsnIleTyrIleValAlaGlyValPheIleAlaIleSerLeuGlnI 248
DB 66 CTCGCCGGTAACTTAAATGTGGCTGGCTTCATCGGCATCTGCTTTCAGAT 125
QY 248 ePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlnAlaVal-LysAlaGlyH 268
DB 126 ATTTGGCATCTTCCTGGCAGTACGCTGATCTCAGACATCGAGCAGTGAAGCGCGGCC 185
QY 268 IShiSpHe 270
DB 186 ATCACTTC 193

RESULT 11
US-09-895-828-430
; Sequence 430, Application US/09895828
; Patent No. US20020099012A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Carter, Darrick
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.539
; CURRENT APPLICATION NUMBER: US/09/895,828
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11, 662
; OTHER INFORMATION: n = A,T,C or G
US-09-895-828-430

Alignment Scores:
Pred. No.: 3.88e-25 Length: 690
```

```

Score: 282.00 Matches: 58
Percent Similarity: 50.25% Conservative: 43
Best Local Similarity: 28.86% Mismatches: 78
Query Match: 19.46% Indels: 22
DB: 10 Gaps: 3

US-09-905-674-2 (1-270) x US-09-895-828-430 (1-690)
QY 63 ValIleuValIleuMetValGlyValValIlePheThrLeuGlyPheAlaGlyCysValGly 82
DB 51 GTGGTCATCATCGCAGTGGGTGCTCTTCTCTGTTGGCTTTGTGGGCTGCTGGGG 110
QY 83 AlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIlePhe 102
DB 111 GCGTGCAAGGAAACCTATTGCTTATGATACAGCTTGCCATCTTCTGCTCTTATCATG 170
QY 103 PheLeuGluLeuAlaValAlaValIleuAlaPheLeuPheGlnAspTrpValArgAspArg 122
DB 171 TTGGTGAGGTGGCGCGCATTCGCTGCTATGTGTTAGATAGATAGATGATGTCAGAG 230
QY 123 PheArgGluPhePheGluSerAsnIleLysSerTyrArgAspAspIleAspLeuGlnAsn 142
DB 231 TTATATTAACAACCTCCGGCAGCAGATGAGATTAACCGAATAACAACCACTGCTTCG 290
QY 143 LeuIleAspSerLeuGlnLysAlaAsnGlnCysCysGlyAlaTyrGlyProGluAspTrp 162
DB 291 ATCTCGGACAGATGACAGCAGATTTTAAGTCGTGGGTGCTACTACACAGATTTG 350
QY 163 AspleuAsnValTyrPheAsnCysSerGlyAlaSerTyrSerArgGlyLysCysGlyVal 182
DB 351 GAG-----AAAATCCCTTCACATGTCGAAACCGAGTC 383
QY 183 ProPheSerCysValProAspProAlaGlnLysValValAsnThrGlnCysGlyTyr 202
DB 384 CCCGACTCTCTGTCATTAAT-----GTTACTGTGGCTGGGATT 425
QY 203 AspValArgIleGlnLeuLysSerLysTrpAspLeuSerIlePheThrLysGlyCysIle 222
DB 426 AATTTCAC-----GAGAAAGCATTCATTAAGAGGGCTGCTG 464
QY 223 GlnAlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAla 242
DB 465 GAGAAAGATTGGGGCTGCTGAGGAAATGTGCTGTGAGCTGACGACAGCCCTTGA 524
QY 243 IleSerLeuGlnIlePheGlyIlePheLeuAlaArgThrLeuIleSerAspIleGlu 262
DB 525 ATTGCTTTTGTGAGGTTTGGGAATGTCTTCTGCTGCTGCTGGAAGAGTATCAGA 584
QY 263 Ala 263
DB 585 AGT 587

RESULT 12
US-10-042-417-37
; Sequence 37, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-15
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
```



```

; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN MEMBRANE SPANNING PROTEINS
; FILE REFERENCE: PF-0489-1 CON
; CURRENT APPLICATION NUMBER: US/09/823,356
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/039,307
; PRIOR FILING DATE: 1998 March 13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PERL Program
; SEQ ID NO: 32
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010025098A1 2779610
US-09-823-356-32

```

```

Alignment Scores:
Pred. No.: 2,63e-23 Length: 2358
Score: 273.50 Matches: 67
Percent Similarity: 44.48% Conservative: 46
Best Local Similarity: 26.38% Mismatches: 100
Query Match: 18.88% Indels: 41
DB: Gaps: 8

```

US-09-905-674-2 (1-270) x US-09-823-356-32 (1-2358)

```

QY 11 ValSerCysTrpTyrIleuLeuPheSerTyrAsnIleIlePheThrLeuAlaGly 30
DB 156 ATTACTGT---TTCAGAGCGCTTCTGTAATCTACACTTTATTTCTGATCATGCG 212
QY 31 ValValPheLeuGlyValGlyLeuTrpAlaTrpSerGluGlyValLeuSerIleu 50
DB 213 GTTATCTCTTTCGACGTTGGCATTTGGGCG-----AAGGTGAGCGCTGAGATAC 263
QY 51 ThrIysValThrArgMetHisGlyIleAspProValIleuValIleuMetValGlyAl 70
DB 264 TTTTCTCTTTAAATGAGAGGCCACCAATGCCCCCTGCTGCTACTGCTACTGCTAC 323
QY 71 ValMetPheThrLeuGlyPheAlaGlyCysValGlyAlaLeuArgGluAsnIleCysLeu 90
DB 324 GTCATTATCTTCTTGGCGACCTTGTGTTGCTTACTGCTGACCTGCTGATGATG 383
QY 91 LeuAsnPhePheCysGlyThrIleValIleuIlePhePheLeuGluAlaValAlaVal 110
DB 384 CTAAACCTGTATGCAATGTTTCTGACTCTGTTTGTGTCGACGCTGCTGCTGATC 443
QY 111 LeuAlaPheLeuPheGlnAspTrpValArgAspArgPheArgGluPhePheGluSerAsn 130
DB 444 GTAGATTTGTTTTCAGACATGAGATTAAAGACACCTTTAAGATAATTATGAGAGGCT 503
QY 131 IleIysSerTyrArgAspIleAspLeuGln---AsnLeuIleAspSerLeuGlnIys 149
DB 504 TTGAAGACATATACCTACAGAGAGATTATAGACCATGACGTAGACAAGATCCAAAT 563
QY 150 AlaAsnGlnCysCysGlyValAlaIleTyrGlyProGluAspTrpAspLeuAsnValTyrPheAsn 169
DB 564 ACCTTGCTGTTGTTGGTGTCCACCATTAATAGAGATTGG----- 602
QY 170 CysSerGlyAlaSerTyrSerArgGluIysCysGlyValProPheSerCys----- 186
DB 603 ---ACAGATACATAATTATTACTCAGAAAA---GGATTTCCTAAGAGTTGCTGTAACCT 656
QY 187 ---CysValProAsp---ProAlaGlnIysValIleAsnThrGlnCysGlyTyrAsp 203
DB 657 GAAGATTGTACTCCACAGAGACATGACAAAGTAACAAAT----- 698
QY 204 ValArgIleGlnLeuIysSerIleTrpAspGluSerIlePheThrIysGlyCysIleGln 223
DB 699 -----GAAGTTGTTTATA 713

```

```

QY 224 AlaLeuGluSerTrpLeuProArgAsnIleTyrIleValAlaGlyValPheIleAlaIle 243
DB 714 AAGGTGATGACCATTTATAGACTCAGAAATGGAGTCTGTCAGGAATTTCTTGGAGTT 773
QY 244 SerLeuLeuGlnIlePheGlyIlePheLeuAlaArgThrLeu 257
DB 774 GCTTGCTTCCAACTGATGTAATCTTCTGCGCTACAGCCCTC 815

```

RESULT 15

US-09-925-301-392

; Sequence 392, Application US/09925301

; Patent No. US20020052308A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA106

; CURRENT APPLICATION NUMBER: US/09/925,301

; PRIOR APPLICATION NUMBER: PCT/US00/05882

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1694

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 392

; LENGTH: 1545

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (24)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (25)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (54)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: misc_feature

; LOCATION: (38)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-925-301-392

Alignment Scores:

Pred. No.: 1.13e-22 Length: 1545

Score: 266.00 Matches: 74

Percent Similarity: 44.48% Conservative: 51

Best Local Similarity: 26.33% Mismatches: 108

Query Match: 18.36% Indels: 48

DB: Gaps: 11

US-09-905-674-2 (1-270) x US-09-925-301-392 (1-1545)

```

QY 6 TyrSerAsnAlaIysValSerCys-----TrpTyrIysTyrIleuLeuPheSerTyr 22
DB 98 TTCAAGACGAGAAAGAAACACATGCGACCGCTTGGCTCAAGTACTGCTTACTTAC 157
QY 23 AsnIleIlePheThrPheLeuAlaGlyValAlaPheLeuGlyValGlyLeuTrpAlaTrpSer 42
DB 158 AATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 217
QY 43 GluIysGly---ValLeuSerAspLeuThrIleValThrArgMetHisGlyIleAspPro 61
DB 218 CTCAAGAGTACATACATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 268
QY 62 ValValLeuValLeuMetValGlyValAlaMetPheThrLeuGlyPheAlaGlyCysVal 81
DB 269 GCCTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 328
QY 82 GlyAlaLeuArgGluAsnIleCysLeuLeuAsnPhePheCysGlyThrIleValIle 101
DB 329 GCCACCTTCAAGAGCGCTCGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388

```

